

ajer

**THE ALBERTA JOURNAL OF
EDUCATIONAL RESEARCH**

VOLUME XLI

NUMBER 4

DECEMBER 1995

The Alberta Journal of Educational Research

Published in March, June, September, and December by the Faculty of Education, University of Alberta.

AJER is a quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Judy Cameron

Managing Editor: Naomi Stinson

Proofreader: Karen McFarlane

Translator: Ronald N. Boivin

Consulting Editors

Terry Belke

Mount Allison University

George Buck

University of Alberta

Ardra L. Cole

*Ontario Institute for Studies
in Education*

John Connors

Canadian Union College

Samuel Deitz

Georgia State University

Sharon M. Haggerty

University of Western Ontario

Antoinette Oberg

University of Victoria

Ruth Rees

Queen's University

Heather Ryan

Massey University

Robert H. Short

University of Alberta

Kelleen Toohey

Simon Fraser University

Robert Wilson

Queen's University

Faculty Advisory Committee

John G. Paterson (Chair)

Joyce Edwards

Robert H. Short

Fern Snart

Ken Ward

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, the Alberta Advisory Committee for Educational Studies, and the Research Grants Office of the University of Alberta

The subscription rate is \$37.00 per year for individuals, \$45.00 per year for institutions. Add \$8.00 for delivery outside Canada. Single copies are \$8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to:

The Alberta Journal of Educational Research

Office of the Dean

845 Education Centre South

University of Alberta

Edmonton, AB, Canada, T6G 2G5.

Fax: (403) 492-0236

Publications Mail Registration Number 1436



Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XLI, Number 4, December 1995

Judy Cameron **377** From the Editor

Articles

Stephen P. Norris **378** Format Effects on Critical Thinking
Test Performance

José L. da Costa **407** Teacher Collaboration:
A Comparison of Four Strategies

Judith Golec **421** University Performance of
John Gartrell Nonmatriculated Admissions
William J. Sveinson

June Beynon **435** Access and Aspirations: Careers in
Kelleen Toohey Teaching as seen by Canadian University
Students of Chinese and Punjabi-Sikh
Ancestry

Marianne Gareau **462** Parents and Schools Working Together:
Don Sawatzky A Qualitative Study of Parent-School
Collaboration

G. Patrick O'Neill **474** Authorship Patterns in AJER:
Forty Years in the Making

R.S. Brown **482** Manifestation of Latent Thetas:
J.O. Anderson A Comparison of Field-Study and
M.M. McGee Positivistic Approaches to the Investigation
of Self-Nurturance

From The Editor

This issue marks the end of my tenure as Editor. I would like to take this opportunity to thank the many individuals who made my experience as Editor rewarding. Several people deserve recognition for their efforts in producing a top journal. Thanks to all the authors who submitted papers and to the Editorial Board and reviewers who offered their critical and incisive comments. I would also like to express my gratitude to the staff in the Dean's Office who attended to aspects of finance, printing, and mailing, and to the *AJER* Faculty Advisory Committee for their support and encouragement. A special thanks to the managing editor Naomi Stinson for transforming manuscripts into the high quality journal that you are reading now.

It is my pleasure to introduce the new Editor Dr. Beth Young from Educational Policy Studies. Dr. Young brings a new perspective to *AJER*. Her research interests include gender issues in Canadian educational administration and leadership; her focus is on women's careers in education and related policy issues. On behalf of myself and all those involved in *AJER*, we welcome Beth to her new role.

Judy Cameron

Stephen P. Norris

Memorial University of Newfoundland

Format Effects on Critical Thinking Test Performance

The influence of five format differences on the scores of 15 critical thinking measures was examined. The format differences were: whether the measures were in multiple-choice or constructed-response format; whether among the multiple-choice measures all items had a common set of response alternatives, or specific alternatives for each item or subset of items; whether or not examinees were requested to provide justifications for their answers; whether or not examinees were required to make judgments of credibility of information or of people; and whether or not the items were cast in the context of a narrative story. Using these format differences, a variety of construct models were hypothesized to explain correlations among scores on the measures. Confirmatory factor analysis was used to test the ability of each of the hypothesized models to account for the observed correlation matrix calculated from the performances of 172 senior high school students on all the measures. Only one model was able to pass all the tests of goodness of fit applied to the data, including the χ^2 test of perfect fit between the observed correlations and the correlations implied by the hypothesized model. That model postulated five factors: a common critical thinking factor, a narrative context factor, a nonnarrative context factor, a credibility judgment required factor, and a no credibility judgment required factor. One of the most significant conclusions from this finding is that much of the debate over the relative merits of multiple-choice and constructed-response critical thinking tests is ill founded: these two formats appear to be equivalent.

Cette étude examine l'influence de cinq différentes sortes ou formats de testing sur les scores de 15 moyens d'évaluation de la pensée critique. Les différences entre les sortes ou formats de tests étaient: si les réponses étaient du genre choix multiples ou réponses à développement; si parmi les réponses à choix multiples il existait un ensemble commun de réponses alternatives possibles, ou des alternatives spécifiques pour chaque item ou chaque sous groupes d'items; si on demandait ou ne demandait pas à chaque candidat(e) à l'examen de justifier ses réponses; si les candidat(e)s devaient ou ne devaient pas porter jugement sur la crédibilité de l'information ou des personnages présentés; et si les items étaient présentés ou pas dans le contexte d'une histoire au style narratif. C'est en utilisant ces différents formats et ces différentes sortes d'examens présentés qu'on a pu construire une variété de modèles pour expliquer la variation et la covariation des scores selon les différents moyens d'évaluation évalués. On utilisa l'analyse des facteurs de confirmation pour vérifier l'habileté de chaque moyen d'évaluation sur les facteurs hypothétiques pour expliquer la matrice de variation et de co-variation observées et calculées d'après les performances de 172 étudiant(e)s du niveau secondaire 2^e cycle sur les 15 moyens d'évaluation. Seulement un modèle semblait remplir les critères de satisfaction appliqués aux données y inclus le test χ^2 d'exactitude entre les corrélations observées et les corrélations sous-entendues par le modèle hypothétique. Ce modèle a postulé cinq différents facteurs: un facteur de pensée critique commun, un facteur de contexte narratif, un facteur de contexte non-narratif, un facteur qui nécessiterait un jugement de crédibilité, et un facteur qui requierrait un jugement de non-crédibilité. Une des conclusions les plus significatives de cette recherche est que les mérites relatifs sur le débat des questions à choix multiples contre celles des questions à développement est en grande partie

Stephen Norris is a professor of educational research and philosophy of education in the Faculty of Education. His areas of research interest include the philosophy and policy of science education, philosophy of reading, reading science text, and critical thinking testing.

réponses mal fondé. Ces deux genres de questions, celles à choix multiples et celles à développement, semblent s'équivaloir.

Do differences in format affect what tests of critical thinking measure? This question generates much opinion and controversy, but there is insufficient evidence to answer it adequately. In this study the influence on scores of several format differences was examined for 15 measures of critical thinking.

The format difference that perhaps has caused most of the debate over critical thinking testing is between multiple-choice and constructed-response tests. Research on the relationship between multiple-choice and constructed-response tests began in the 1920s. According to Hogan (1981), who conducted an extensive review of studies up to 1981, the chief concern in the earliest investigations was the comparative reliability of the two formats. However, the concerns changed to include the question of equivalence. Hogan believed at the time of writing his report that regardless of the research "popular opinion on this question [was] rather well formulated and almost universally negative, i.e., the two types of items do *not* measure the same thing" (p. 2). This opinion is especially significant in the area of critical thinking testing. One argument used to support the opinion is that tests of subject matter that employ the multiple-choice format may be able to test effectively for knowledge recall, but it requires constructed-response tests to assess ability to think at higher levels about the content. Because critical thinking tests are by definition tests of higher level thinking, the conclusion is obvious: the term *multiple-choice critical thinking test* is an oxymoron. Hence the issue is salient in the domain of critical thinking as it is contended that as a matter of principle *nothing* in the domain is testable using multiple-choice tests. If this contention is true, then all critical thinking testing should follow the more time-consuming and expensive (by a factor of about 1,000, according to Ennis & Norris, 1990) route of constructed-response formats.

In his review, Hogan (1981) drew a series of conclusions based on three types of studies: direct correlation with correction for attenuation of multiple-choice and constructed-response tests to check for divergence of the correlation coefficient from unity; correlation of multiple-choice and constructed-response tests with a separate criterion measure to check which format yields the higher correlation; and treatment application to determine which type of measure is more sensitive to detection of a predicted change. His generalizations that are most relevant to this study are: (a) in most instances, multiple-choice and constructed-response measures were found to be equivalent or nearly so; and (b) multiple-choice tests tend to correlate more highly than constructed-response tests with external criteria.

In a more recent review Traub and MacRury (1990) drew different conclusions. Acknowledging that the differences are not well understood, they concluded that the two formats appear to measure somewhat different abilities, although any difference that exists is slight, if the constructed-response format requires answers no longer than a sentence or two for each item. Traub (1993) further argued that differences between the formats should not be expected if items require the manipulation of data and ideas. Barnett-Foster and Nagy (1995) found "the two formats are remarkably similar for first-year

university students in science and science-related programs of study" (p. 33). However, they cautioned that the conclusion should not be extended to tests of other types of reasoning strategies.

Thus there is sufficient uncertainty about the relationship between multiple-choice and constructed-response formats to warrant further research in the area. There is additional motivation to explore the issue using critical thinking tests, because they have not been studied previously with regard to this topic and because the opinion of many is that in principle critical thinking can be tested only with constructed-response formats. Even if the multiple-choice versus constructed-response difference makes no difference to critical thinking testing, it is still worthwhile exploring the influence of other differences in format.

Included among the 15 measures examined in this study were six in constructed-response format, and nine in multiple-choice format. In addition to variation along the multiple-choice/constructed-response format dimension, the 15 measures varied along several other dimensions that could affect scores. First, in the multiple-choice format four measures used a common set of alternatives for all items, and five measures used a specific set of alternatives for each item or each subset of items. Second, in the constructed-response format three measures requested examinees to justify their answers with reasons, and three measures did not request justification. Third, seven measures (4 in multiple-choice format and 3 in constructed-response format) asked examinees to judge the credibility of information presented; the other eight measures did not ask for such judgments. Finally, eight measures (5 multiple-choice and 3 constructed-response) embedded items in the context of story narratives; the other seven measures used either expository or descriptive portrayals of content for contextualizing items.

Based on these five format differences (multiple-choice vs. constructed-response; common-alternatives vs. specific-alternatives; justification-requested vs. justification-not-requested; credibility-judgment-required vs. credibility-judgment-not-required; narrative-context vs. nonnarrative-context), a variety of construct models were hypothesized to explain variation and covariation among scores. Confirmatory factor analysis was used to test the ability of each hypothesis to account for the observed correlation matrix among scores on the 15 measures.

More specifically, the construct models displayed in Table 1 were hypothesized. The models were based on combinations of the 15 factors listed in the left-hand column. The first group of six factors are simple factors because each refers only to one format characteristic. The second group of eight factors are compound factors because each refers to two format characteristics. Finally, a common critical thinking factor was used in building some models. Using various combinations of these 15 factors, 12 construct models were built: three each of three-factor, four-factor, five-factor, and six-factor models. An attempt was made to test each model under two assumptions: (a) with all factors constrained to be orthogonal, and (b) with at least some of the format factors correlated. In two instances, however, models with correlated factors were insufficiently constrained to allow estimation of the models to proceed.

Table 1
Hypothesized Construct Models to Explain Correlations Among 15 Measures

Factors	Three-Factor Models			Four-Factor Models			Five-Factor Models			Six-Factor Models		
	3A	3B	3C	4A	4B	4C	5A	5B	5C	6A	6B	6C
Simple factors												
multiple-choice (m-c)	X									X		
constructed-response (c-r)	X									X		
narrative context		X		X			X			X		
nonnarrative context		X		X			X			X		
credibility judgment required			X	X			X			X		X
credibility judgment not required			X	X			X			X		X
Compound factors												
m-c/common alternatives					X			X			X	
m-c/specific alternatives					X			X			X	
c-r/justification requested					X			X			X	
c-r/justification not requested					X			X			X	
m-c/narrative						X			X			X
m-c/nonnarrative						X			X			X
c-r/narrative						X			X			X
c-r/nonnarrative						X			X			X
Common critical thinking factor	X	X	X				X	X	X			

The three-factor models all included a common critical thinking factor. In each model the other two factors represented a format difference that divided the sample of measures into two subgroups. Two of the four-factor models (4B and 4C) divided the measures into multiple-choice and constructed-response categories, and in addition used other format factors to subdivide within each of these two main categories. Model 4A referred to format factors that cut across the multiple-choice and constructed-response categories. The five-factor models were constructed by adding a common critical thinking factor to each of the four-factor models. Finally, the six-factor models comprised combinations of the four-factor and three-factor models less the common critical thinking factor. It was not fruitful to combine the compound factors used in model 6B with the credibility judgment factors because the divisions m-c/ common-alternatives, m-c/specific-alternatives, c-r/justification-requested, and c-r/justification-not-requested were coextensive, respectively, with the categories m-c/credibility-judgment-required, m-c/credibility-judgment-not-required, c-r/credibility-judgment-required, and c-r/credibility-judgment-not-required.

Significance

The questions raised in this study have broad educational significance. Critical thinking has been advocated as an essential goal of education throughout this century (Black, 1946; Dewey, 1933; Ennis, 1962, 1981; Hullfish & Smith, 1968; Siegel, 1988; Smith, 1953). However, as Black (1946) said, "To be in a position to improve reasoning means to be in a position to distinguish good reasoning from bad" (p. 7).

Critical thinking is widely taken to be reasonable and reflective thinking that is focused on deciding what to believe or do (Norris & Ennis, 1989), a conception that has evolved from an earlier one by Ennis (1962), which took critical thinking to be the correct assessing of statements. So conceived, the critical thinker has both abilities and dispositions (Ennis, 1981; Perkins, Jay, & Tishman, 1993; Siegel, 1988). The abilities include the ability to maintain clarity, to provide and judge the basic support for actions and beliefs, and to make and judge inferences from basic information (Norris & Ennis, 1989). Dispositions to think critically are also needed, because people can have abilities they do not use. Critical thinkers are disposed, among other things, to seek reasons, consider alternatives, and withhold judgment when the evidence is not sufficient. This study focuses only on tests that are cast as measures of critical thinking abilities. Dispositions possibly play a role in determining performance on these tests, and possibly test format affects the size of that role. Nevertheless, disposition effects are not examined here.

As pointed out in several places (Bennett & Ward, 1993; Ennis & Norris, 1990), there has been a great deal of rhetoric about the adequacy of various formats for critical thinking testing. Some of this rhetoric has issued definitive conclusions about the relative merits of, for example, multiple-choice and constructed-response formats based on questionable claims about one format or the other such as: that scoring constructed-response tests is unreliable, that multiple-choice tests are more objective, that constructed-response tests give better evidence of thinking, that constructed-response tests assess productive and organizational skills, that multiple-choice tests cannot measure higher order thinking, and that constructed-response tests better resemble criterion

behaviors. Although these claims are often presupposed in recommendations for critical thinking testing, little evidence is available to support them. For instance, constructed-response formats are often proposed as a solution to perceived problems with multiple-choice formats on the grounds that constructed responses provide more direct and complete evidence on thinking than multiple-choice answers. However, this rationale relies on unsupported assumptions about the relative validity of different critical thinking test formats and unnecessary assumptions about how multiple-choice tests must be designed (Norris, 1992). Perhaps the two formats measure different constructs; perhaps they measure the same constructs. Perhaps well-designed multiple-choice tests could support the same inferences about critical thinking as are supported by constructed-response tests. Evidence on these and other issues regarding critical thinking testing is needed badly.

It is easy to agree with Resnick (1987) that critical thinking testing has not kept up with the demand for instruction in the area. There are no adequate answers to the questions raised above, despite their recognized importance (Arter & Salmon, 1987; Ennis, 1984; Ennis & Norris, 1990). Virtually all the evidence on critical thinking tests comes from factor analyses of superseded versions of the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) and the Cornell Critical Thinking Tests (Ennis & Millman, 1985). Also many criticisms of such widely used critical thinking tests are based on inadequate evidence. For example, McPeck (1981) concluded only from first-order correlations of 0.55 to 0.75 between the Watson-Glaser test and IQ and reading tests that the Watson-Glaser measures primarily IQ and/or reading ability; and only from inspection that the induction section of the Cornell test Level Z measures mostly reading comprehension. Whimbey (1985) argued that various standardized achievement tests can serve as critical thinking tests, based on correlations between the two types of tests. For example, he reports correlations of .40 to .68 between the New Jersey College Basic Skills Placement Test and the Cornell tests. However, he gives no source for the data and no account of the 55% to 84% of nonoverlapping variance between the tests. This study provides evidence for evaluating criticisms like McPeck's and Whimbey's.

The inability properly to settle these issues given current evidence impinges adversely on educational practice. The critical thinking movement in many countries is motivated in part by a recognition that children learn large amounts of information, but learn less well how to evaluate and use that information. However, unless we understand the techniques for critical thinking assessment and the ways in which critical thinking tests work, we will not be able to determine whether the effort put into critical thinking instruction is leading to desired results. In addition, unless there is evidence on the validity of different critical thinking test formats, then policy decisions about what formats to use will not be grounded adequately.

Method

Definitions of Formats

Multiple choice versus constructed response. Traub's (1993) definitions were used to make the multiple-choice versus constructed-response format distinction: a multiple-choice test contains items "in which the examinee is required to choose an answer from a relatively small set of response options" (p. 29); a

constructed-response test “requires the examinee to compose an answer ... [ranging in length from] a word or a phrase, a number or a formula ... [to] a paragraph, an extended essay, a multistep solution to a mathematical or scientific problem” (pp. 29-30). All the multiple-choice measures used in this study were scored on the basis of a single correct answer for each item; the constructed-response measures contained items requiring from single-word responses to paragraph-length responses.

Among the 15 measures, were five pairs of what Traub and MacRury (1990) called “stem-equivalent” tests. A pair of multiple-choice and constructed-response tests is stem-equivalent if the tests are equivalent item-stem for item-stem, or “when, item for item, the tasks posed in each test are identical except for the response required” (Traub, 1993, p. 32, fn. #2). The pairs of measures were not, however, “scoring-equivalent” (Traub & MacRury, 1990). In multiple-choice testing, items usually are scored dichotomously as either correct or incorrect. All multiple-choice measures used in this study were scored this way, although three measures allowed for “scaled-answer” (Phillips & Patterson, 1987) scoring in which varying credit is assigned depending on the answer selected. In constructed-response testing, partial credit is often given for responses that are considered less than ideal, but somewhat meritorious. Among the five pairs of stem-equivalent measures, three of the constructed-response measures were scored so as to allow for zero, partial, and full credit. The other two constructed-response measures were scored dichotomously. However, the dichotomously scored constructed-response measures were still not scoring-equivalent to their stem-equivalent multiple-choice counterparts, because in each format different judgments were made to determine correctness. In the multiple-choice format, scores reflected whether or not examinees answered questions correctly. In the constructed-response versions, in addition to answering the same questions posed on the multiple-choice versions, examinees were required to justify their answers. Scores reflected whether or not their justifications were sound.

It makes sense to use different approaches and criteria in order to score stem-equivalent multiple-choice and constructed-response tests, because the formats are intended to provide different information for the examiner. Whether that difference in intent translates into measurement differences such that the two formats measure different psychological constructs is an issue to be studied here. It would bias unduly the outcome of the study, however, to adjust the scoring so that they are more similar than they would be in actual use. Hence scoring equivalence was not imposed on stem-equivalent pairs.

Narrative context versus nonnarrative context. The narrative and nonnarrative measures were differentiated, because different text types make distinct demands on readers’ knowledge and expectations about reading, and consequently on their understanding of text. Narration is the most familiar to school students of all text types. It informs readers of what is happening by giving an account of events or actions. Typically, narrative text includes characters, plot, theme, and style. The narrative-context critical thinking tests studied here contained characters and themes; the nonnarrative tests did not.

Credibility judgment required versus no credibility judgment required. Credibility-judgment measures request that examinees decide on the believability of a

statement or of a person. This task is different from judging truth in that a statement may be believable but not true, or true but not believable. In credibility-judgment measures examinees often are presented pairs of statements. Each statement in a pair might be true, but examinees are asked to decide which is more credible. Such decisions are central to critical thinking and hence to many measures of the construct.

Common alternatives versus specific alternatives. The common-alternative versus specific-alternative distinction applies only to the multiple-choice measures. For common-alternative measures the same set of alternatives is used for all items. Such tests are quite prevalent in critical thinking testing. For instance, examinees may be given the choice for each item to judge whether evidence supports a given conclusion, goes against the conclusion, or is irrelevant to the conclusion. By comparison, in the specific-alternative measures examinees must comprehend and take on a new critical thinking task for each item or subset of items.

Justification requested versus no justification requested. The difference between the justification and no-justification measures resides in the former making a specific request for examinees to provide justification for their responses. Having to make their reasons explicit suggests a different and higher cognitive load, although examinees presumably consider justifications while taking the measures that request no justification.

Instruments

Fifteen critical thinking measures were used. Among these measures the following three comprised pairs of stem-equivalent multiple-choice and constructed-response formats: (a) Test on Appraising Observations, Part A; (b) Cornell Critical Thinking Test, Level X, Section I; and (c) Test of Inference Ability in Reading Comprehension. Abbreviated names for all measures and their format characteristics are provided in the first and second columns of Table 3.

Test on Appraising Observations (3 measures). This test contains questions set in the contexts of story narratives and assesses ability to judge the credibility of reports of observations. The multiple-choice format (Norris & King, 1983) consists of 50 items in two parts. Part A has 28 items set in the context of a traffic accident; Part B has 22 items set in the context of a river exploration. Each part was used as a separate measure in this study. Each item provides two statements spoken by characters who witnessed or who were involved in the accident or who made observations along the river. Examinees are to choose which, if either, of the statements they have more reason to believe at the time the statements are made. There are thus three alternatives for each item: (a) the first statement is more believable, (b) the second statement is more believable, or (c) neither statement is more believable, they are equally believable. The constructed-response format (Norris, 1986) is based on Part A of the multiple-choice format. It contains 25 items, and for each item examinees are to judge which, if either, of two statements is more believable. In addition, they are asked to justify their decision. Scores for each item represent whether or not the justification was sound. Some of the relevant factors in judging the statements are: the observer's expertise, alertness, and conflict of interest; the observation conditions; and the source of the observation and the statement reporting it.

Cornell Critical Thinking Test, Level X (2 measures). This test (Ennis & Millman, 1985) has four subsections, with 71 items. Items are cast in the context of a story about a group of scientific explorers from Earth who have gone to the planet Nicoma to search for some lost explorers. Section I, on induction, is the only one used in this study. It contains 23 items dealing with a health officer's hypothesis that the previous explorers are all dead. Different explorers report observations, and in the multiple-choice version examinees are to decide whether each observation report: (a) supports the health officer's idea that everyone in the first group is dead, (b) goes against the health officer's idea, or (c) does not help to decide on the hypothesis. A constructed-response version of Section I was designed for this study by altering the directions to request examinees to tell why they chose their answers and by providing an answer sheet to accommodate their justifications. The score for each item indicates whether or not the justification for the answer-choice is sound.

Test of Inference Ability in Reading Comprehension (6 measures). This test is designed to examine inference ability in the context of reading. It is a 36-item test in both multiple-choice (Phillips & Patterson, 1987) and constructed-response (Phillips, 1989) formats. Each format consists of three subtests based on either narrative, expository, or descriptive text. The first subtest in each format, *UFOs*, is an exposition about unusual atmospheric phenomena; *Money* is a description of the everyday use of money, of how it works, and of its historical development; and *The Wrong Newspapers* is a narrative about a mix-up in newspaper delivery. Each subtest consists of four to five paragraphs with questions after each for a total of 12 questions per subtest. In the multiple-choice format there are four alternatives per item. The constructed-response format uses the same 36 items; students are to formulate their own responses and justify them. Justifications are scored on a scale of 0 to 3 for each item.

Ennis-Weir Critical Thinking Essay Test. The Ennis-Weir (Ennis & Weir, 1985) is intended to evaluate ability to appraise an argument and to formulate a written argument in response. The test begins by asking examinees to read a letter to the editor of a fictional newspaper. In the letter, a proposal is made to end overnight parking on city streets, and a variety of arguments are offered in support of the proposal. Examinees are asked to write a letter evaluating the arguments in each paragraph and in the letter as a whole. The scoring system for each paragraph response is as follows: -1 for judging an argument incorrectly and/or showing bad judgment in justifying; 0 if no response if made; +1 if the argument is judged correctly but it is not justified; +2 if justified semiadequately; and +3 if justified adequately.

Test of Inductive Reasoning Principles. This test (Norris & Ryan, 1992) has 22 items written in the context of reports from biologists, language experts, physicians, and sociologists who are exploring the fictional planet Ladelin. In each question some data and either a generalization from or an explanation of the data is given by two different explorers. Examinees must choose which explorer, if either, gives stronger support for his or her conclusion.

Watson-Glaser Critical Thinking Appraisal. The Watson-Glaser (Watson & Glaser, 1980) comprises five subtests, each consisting of 16 items and covering the following aspects of critical thinking: inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. In subtest I, the

ability to judge inferences made from a statement of facts is tested. There are five judgments from which to choose: true, probably true, insufficient data, probably false, or false. Subtest II provides five statements followed by a number of proposed assumptions. Examinees are to judge whether a person making a given statement is also making the proposed assumptions. Each assumption is to be judged independently from the others. In subtest III, examinees are to judge whether conclusions follow necessarily from given statements. Subtest IV tests for interpretation abilities. It contains five short paragraphs each followed by several suggested conclusions. Examinees are to judge whether each of the proposed conclusions follows beyond a reasonable doubt from the information. Finally, in subtest V, examinees are to judge whether given arguments are strong or weak.

Norris-Ryan Argument Analysis Test. This instrument (Norris & Ryan, 1989) is a 40-item multiple-choice test in five parts testing the ability to identify and distinguish arguments, reasons, and conclusions. Each section consists of a number of short paragraphs followed by questions with four alternative answers.

Participants

Participants were selected from grades 10, 11, and 12 in a large senior high school in central St. John's, Newfoundland. Participation was dependent on written consent from both students and at least one of their parents or guardians. A letter informing students about the study and what would be required if they chose to volunteer was distributed by the school principal to four classes at each grade level. The number of volunteers was 181, but complete data were collected on 172 students, distributed by grade and sex as in Table 2. Students were offered an honorarium for their time and effort. This sample represented a broad range of student abilities, with average school grades ranging from failing to 98%.

Procedure

The testing schedule was as follows: time was reserved each day after school for two weeks; each afternoon one test was given, sometimes including subtests; students were tested as a group in the school's cafeteria; at the end of the two weeks, make-up days were scheduled for students who had missed any testing. Following the recommendations of Traub and Fisher (1977) and Heim and Watts (1967) for the paired stem-equivalent multiple-choice and con-

Table 2
Subjects by Grade Level and Sex

Sex	Grade			Totals
	10	11	12	
Male	28	22	25	75
Female	43	29	25	97
Totals	71	51	50	172

structed-response measures, the constructed-response version of each pair was administered first.

The multiple-choice measures were machine scored using the answer keys and scoring procedures described in the manuals for each measure, with the exception of the multiple-choice version of the Reading Inference measures. For the Reading Inference measures, items were scored initially using two procedures, and the results compared: first, on the basis of the scale described in the manual, which rates answer choices on a range of quality; and, second, as right or wrong on the basis of whether or not the best answer was selected. The (attenuated) correlation between scores graded by both methods was 0.96. The scores from the right-wrong scoring procedure were used for this study on the assumption that it yielded the same information as the more complicated scaled-answer procedure.

For the constructed-response measures, the scoring systems described in the manuals were used, except for the Cornell Level X. The constructed-response format for this measure was developed for this study, so a scoring guide also was developed based on the justifications offered in the test manual (Ennis, Millman, & Tomko, 1985) for the keyed answers to each multiple-choice item.

Analysis

Descriptive statistics and a correlation matrix were computed for the 15 measures. A principal components analysis was performed and eigenvalues were examined to help determine a suitable number of factors. The viability of the various hypothetical factor models to account for the correlations among the measures was examined using EzPATH, Version 1.0 (Steiger, 1989), an implementation of structural equation modeling used by SYSTAT. Using the correlation matrix as input, EzPATH was directed to compute factor loadings with no constraints on the sizes or signs of the loadings. For each model EzPATH also was directed to compute a unique factor loading for each measure, again with the sizes and signs of the loadings unconstrained. For each model, solutions for both orthogonal and oblique common factors were attempted. The method of estimation first produced least squares estimates for factor loadings and correlations, and then used these estimates as input to a maximum likelihood estimation. In two cases it proved impossible to estimate the models because of insufficient constraints that led to out-of-bounds estimates of some factor loadings (loadings exceeding 1.0).

Three indices were used to judge the goodness of fit between the observed correlation matrix used as input and that computed from the estimated model. The null hypothesis of perfect match between the matrices was tested using the chi-squared statistic. The probability levels reported are the probabilities of obtaining values of χ^2 equal to or greater than the obtained values on the hypothesis that the computed matrices perfectly fit the observed matrix.

In addition to the test of perfect fit supplied by the chi-squared, degree of goodness of fit was measured by two indices: the Steiger-Lind RMS Index (R^*) and the Adjusted Population Gamma Index (Γ_2). Values of R^* less than .10 indicate a reasonably good fit, values below .05 indicate an excellent fit, and values less than .01 indicate an outstanding fit. Values of Γ_2 greater than .90 indicate a good fit and values greater than .95 indicate an excellent fit. The 90%

Table 3
Format Characteristics, Numbers of Items, Means, Standard Deviations, Alpha Reliabilities,
and Correlation Matrix for All Measures

Measure	Format	No. of items	Mean	S.D.	Alpha	No. of measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Observation Appraisal (A)	CR ¹ /J ³ /NARR ⁷ /CRED ⁹	25	16.0	3.86	.71	1	1.00														
Cornell, Level X (I)	CR/J/NARR/CRED	23	15.4	3.79	.72	2	.335	1.00													
Reading Inference UFOs	CR/NJ ⁴ /NONNARR ⁸ /NOCRED ¹⁰	12	19.1	4.16	.43	3	.280	.277	1.00												
Reading Inference Money	CR/NJ/NONNARR/NOCRED	12	17.7	5.20	.58	4	.383	.209	.536	1.00											
Reading Inference News	CR/NJ/NARR/NOCRED	12	20.8	5.32	.59	5	.337	.205	.390	.294	1.00										
Emmis-Weir	CR/J/NONNARR/CRED	9	5.15	6.30	.72	6	.250	.323	.250	.310	.305	1.00									
Inductive Reasoning	MC ² /CA ⁵ /NARR/CRED	22	10.6	3.15	.52	7	.309	.228	.174	.192	.210	.144	1.00								
Reading Inference UFOs	MC/SA ⁶ /NONNARR/NOCRED	12	8.29	2.16	.63	8	.360	.302	.314	.302	.316	.308	.314	1.00							
Reading Inference Money	MC/SA/NONNARR/NOCRED	12	7.71	1.70	.37	9	.220	.205	.145	.268	.255	.164	.128	.280	1.00						
Reading Inference News	MC/SA/NARR/NOCRED	12	7.87	2.02	.47	10	.268	.282	.240	.300	.344	.268	.224	.428	.417	1.00					
Watson-Glaser	MC/SA/NONNARR/NOCRED	80	48.6	8.06	.75	11	.257	.292	.377	.477	.342	.276	.258	.486	.343	.307	1.00				
Observation Appraisal (A)	MC/CA/NARR/CRED	28	18.5	3.30	.54	12	.584	.338	.359	.393	.246	.185	.312	.350	.359	.395	.290	1.00			
Observation Appraisal (B)	MC/CA/NARR/CRED	22	11.8	3.01	.52	13	.342	.291	.236	.285	.055	.183	.355	.320	.254	.234	.332	.565	1.00		
Argument Analysis	MC/SA/NONNARR/NOCRED	40	20.4	5.55	.76	14	.410	.352	.437	.525	.352	.407	.398	.508	.393	.388	.549	.461	.409	1.00	
Cornell, Level X (I)	MC/CA/NARR/CRED	23	15.9	3.14	.67	15	.288	.567	.222	.225	.276	.319	.342	.315	.242	.288	.301	.295	.303	.386	1.00

1 CR = constructed response

2 MC = multiple choice

3 J = justification requested

4 NJ = justification not requested

5 CA = common alternatives for all items

6 SA = specific alternatives for each item or item subset

7 NARR = items set in narrative context

8 NONNARR = items set in nonnarrative context

9 CRED = credibility judgment required

10 NOCRED = credibility judgment not required

confidence intervals on R^* and Γ_2 were also computed in order to judge degree of confidence in the estimates of goodness of fit.

Results

Descriptive Statistics

Table 3 contains the format characteristics, number of items, mean, standard deviation, and alpha reliability for each measure, and the correlations among the measures. The highest reliability estimates, 0.75 for the 80-item Watson-Glaser test and 0.76 for the 40-item Argument Analysis test, are not high by the standards of many psychological tests. However, reported reliabilities for critical thinking tests tend to be low (Ennis & Norris, 1990).

Number of Factors

Table 4 contains the eigenvalues associated with the principal components. In determining the number of factors to extract, a variety of tests were applied to the pattern of eigenvalues, following the advice of Kim and Mueller (1978) who suggest seeking a convergence of a number of tests. The number of eigenvalues greater than one suggests four factors. If accounting for 5% of the variance is taken as a minimum amount for substantive importance of any component, then six factors seem correct. Applying the scree test suggests six, possibly four, or as few as two factors, depending on how the graph is viewed.

Theoretical consideration of the influence of format on the scores of the critical thinking measures used in this study suggested models with from three to six factors (see Table 1). The empirical data support testing models with either of these numbers of factors. Thus the judgment was made to test all the models proposed in Table 1.

Three-Factor Models

Table 5 presents the factor loadings, and, in the case of correlated factors, the correlations between format factors for Models 3A, 3B, and 3C. The top half of the table presents loadings for the orthogonal solutions, and the bottom half for the oblique solutions.

Almost without exception, measures load heavily on the Common Critical Thinking factor for each model. In the orthogonal solution to Model 3A the only multiple-choice measures to load on the Multiple-choice factor were the two Observation tests and the Induction test; of the constructed-response measures only the Reading measures loaded on the Constructed-response factor. The oblique solution to Model 3A produced much more clearly defined

Table 4
Eigenvalues Associated with Principal Components

Component	Eigenvalue	Component	Eigenvalue	Component	Eigenvalue
1	5.51	6	0.85	11	0.46
2	1.26	7	0.73	12	0.42
3	1.13	8	0.65	13	0.36
4	1.03	9	0.61	14	0.31
5	0.89	10	0.54	15	0.29

factors, although the Ennis-Weir and the Watson-Glaser did not load on the Common Critical Thinking factor.

In Model 3B neither the orthogonal solution nor the oblique solution yielded clearly defined Narrative and Nonnarrative factors. In Model 3C, no clear Credibility Judgment factor emerged in either the orthogonal solution or oblique solution; however, a No Credibility Judgment factor was present in both solutions.

The largest change in loadings between the orthogonal and oblique solutions was for Model 3A, where the loadings for the Multiple-choice and Constructed-response factors tended to be substantially higher for the solution with correlated factors. The Multiple-choice and Constructed-response factors were highly correlated in the oblique solution to Model 3A.

In Models 3B and 3C there was less of a marked difference in loadings between the orthogonal and oblique solutions. This is consistent with the low correlations between the Narrative and Nonnarrative factors and between the Credibility Judgment and No Credibility Judgment factors in the oblique solution.

Table 9 displays the results of the goodness of fit tests. In each case, the value of the χ^2 leads to a rejection of the hypothesis of perfect fit between the observed and computed correlation matrices. Examining R^* indicates values between .05 and .10 for each model, indicating reasonably good fit. In each test the upper bound of the 90% confidence interval on R^* was less than .10 increasing confidence in the judgment of a reasonably good fit. The Γ_2 values all are greater than .90, indicating good fit. However, in only one case is the lower bound of the 90% confidence interval greater than .90, reducing the confidence in the judgment of good fit for five of the models. Both the R^* and Γ_2 indices indicate that the orthogonal solution to Model 3C fits the data best in this group. However, none of the models fits the data excellently.

Four-Factor Models

Table 6 presents the factor loadings, and in the case of correlated factors the correlations between format factors for Models 4A, 4B, and 4C.

In the orthogonal solution to Model 4A, measures tend to load heavily on their respective factors. In the Narrative and Nonnarrative factors, however, three measures do not load. An oblique solution to Model 4A with all correlations among factors unconstrained could not be tested without finding out-of-bounds estimates for some factor loadings. Attempts to solve this problem by placing further constraints on correlations among factors were made until the solution reported was found. The oblique solution is less easily interpreted than the orthogonal solution because of the negative loadings in the Narrative and Nonnarrative factors. The Credibility Judgment and No Credibility Judgment factors contain loadings very similar to the orthogonal solution.

In both Models 4B and 4C both the orthogonal and the oblique solutions yielded four clearly defined factors. In the oblique solutions the factors were highly intercorrelated, and there was no appreciable change in factor loadings over the orthogonal solutions.

Table 9 displays the results of the goodness of fit tests. In each case, the value of the χ^2 leads to a rejection of the hypothesis of perfect fit between the observed and computed correlation matrices. Examining R^* and its confidence

Table 5
Factor Loadings for Three-Factor Models

Measure: Format	Model 3A Factors			Model 3B Factors			Model 3C Factors		
	(1) Common Critical Thinking	(2) Multiple- Choice Format	(3) Constructed- Response Format	(1) Common Critical Thinking	(2) Narrative Context	(3) Non- Narrative Context	(1) Common Critical Thinking	(2) Credibility Judgment	(3) No Credibility Judgment
Orthogonal Solutions									
Observation A: CR/U/NARR/CRED	.58	*	-.02	.53	.34	*	.62	.21	*
Cornell: CR/U/NARR/CRED	.52	*	.02	.54	.04	*	.62	-.30	*
Reading UFOs: CR/U/NO/NARR/NO/CRED	.51	*	.86	.49	*	.37	.42	*	.42
Reading Money: CR/U/NO/NARR/NO/CRED	.59	*	.27	.51	*	.76	.46	*	.49
Reading News: CR/U/NO/NARR/NO/CRED	.50	*	.16	.51	-.05	*	.38	*	.33
Ennis-Weir: CR/U/NO/NARR/NO/CRED	.48	*	.01	.48	*	.08	.43	-.24	*
Induction: MC/CA/NARR/CRED	.44	.17	*	.47	.06	*	.47	-.04	*
Reading UFOs: MC/SA/NO/NARR/NO/CRED	.64	.03	*	.66	*	-.04	.51	*	.35
Reading Money: MC/SA/NO/NARR/NO/CRED	.48	.04	*	.48	*	.02	.41	*	.23
Reading News: MC/SA/NARR/NO/CRED	.56	-.03	*	.56	.10	*	.49	*	.22
Watson-Glaser: MC/SA/NO/NARR/NO/CRED	.64	.05	*	.63	*	.21	.45	*	.56
Observation A: MC/CA/NARR/CRED	.61	.32	*	.57	.82	*	.76	.50	*
Observation B: MC/CA/NARR/CRED	.46	.89	*	.48	.35	*	.59	.21	*
Argument MC/SA/NO/NARR/NO/CRED	.78	.06	*	.76	*	.18	.64	*	.45
Cornell: MC/CA/NARR/CRED	.53	.07	*	.57	-.03	*	.64	-.40	*

Table 5 (continued)

Measure: Format	Model 3A Factors			Model 3B Factors			Model 3C Factors		
	(1) Common Critical Thinking	(2) Multiple- Choice Format	(3) Constructed- Response Format	(1) Common Critical Thinking	(2) Narrative Context	(3) Non- Narrative Context	(1) Common Critical Thinking	(2) Credibility Judgment	(3) No Credibility Judgment
Oblique Solutions									
Observation A: CR/U/NARR/CRED	.49	*	.43	.51	.38	*	.52	.37	*
Cornell: CR/U/NARR/CRED	.22	*	.45	.55	.07	*	.61	.00	*
Reading UFOs: CR/U/NOONARR/NOCREd	.22	*	.57	.44	*	.44	.38	*	.54
Reading Money: CR/U/NOONARR/NOCREd	.24	*	.61	.45	*	.75	.41	*	.65
Reading News: CR/U/NARR/NOCREd	.11	*	.53	.50	-.01	*	.43	*	.25
Ennis-Weir: CR/U/NOONARR/CRED	.04	*	.53	.48	*	.10	.51	-.11	*
Induction: MC/CA/NARR/CRED	.19	.41	*	.47	.08	*	.48	.07	*
Reading UFOs: MC/SA/NOONARR/NOCREd	.16	.63	*	.66	*	.03	.60	*	.19
Reading Money: MC/SA/NOONARR/NOCREd	.24	.42	*	.46	*	.08	.43	*	.17
Reading News: MC/SA/NARR/NOCREd	.26	.48	*	.55	.13	*	.52	*	.16
Watson-Glaser: MC/SA/NOONARR/NOCREd	.08	.69	*	.61	*	.27	.53	*	.40
Observation A: MC/CA/NARR/CRED	.95	.31	*	.51	.86	*	.54	.82	*
Observation B: MC/CA/NARR/CRED	.47	.37	*	.46	.38	*	.49	.37	*
Argument MC/SA/NOONARR/NOCREd	.23	.76	*	.74	*	.27	.68	*	.38
Cornell: MC/CA/NARR/CRED	.15	.49	*	.59	-.01	*	.66	-.08	*
Correlations Among Factors									
(1)	—			—			—		
(2)	*	—		*	—		*	—	
(3)	*	.88	—	*	.25	—	*	.29	—

*Parameter constrained to be zero.

Table 6
Factor Loadings for Four-Factor Models

Measure: Format	Model 4A Factors				Model 4B Factors				Model 4C Factors			
	(1) Narrative Context	(2) Non- Narrative Context	(3) Credibility Judgment	(4) No Credibility Judgment	(1) Multiple- Choice Common Alternative	(2) Multiple- Choice Specific Alternative	(3) Constructed- Response Justification	(4) Constructed- Response No Justification	(1) Multiple- Choice Narrative	(2) Multiple- Choice Non- Narrative	(3) Constructed- Response Narrative	(4) Constructed Response Non- Narrative
Orthogonal Solutions												
Observation A: CR/U/NARR/CRED	.51	*	.33	*	*	*	.51	*	*	*	.74	*
Cornell: CR/U/NARR/CRED	.17	*	.69	*	*	*	.66	*	*	*	.45	*
Reading UFOs: CR/U/NO/NARR/NOCRE	*	.47	*	.41	*	*	*	.84	*	*	*	.66
Reading Money: CR/U/NO/NARR/NOCRE	*	.72	*	.42	*	*	*	.64	*	*	*	.81
Reading News: CR/U/NARR/NOCRE	.02	*	*	.50	*	*	*	.46	*	*	.45	*
Ennis-Weir: CR/U/NO/NARR/CRED	*	.21	.40	*	*	*	.49	*	*	*	*	.38
Induction: MC/CA/NARR/CRED	.22	*	.38	*	.48	*	*	*	.48	*	*	*
Reading UFOs: MC/SA/NO/NARR/NOCRE	*	.02	*	.69	*	.68	*	*	*	.66	*	*
Reading Money: MC/SA/NO/NARR/NOCRE	*	.03	*	.51	*	.52	*	*	*	.48	*	*
Reading News: MC/SA/NARR/NOCRE	.17	*	*	.52	*	.56	*	*	.46	*	*	*
Watson-Glaser: MC/SA/NO/NARR/NOCRE	*	.25	*	.64	*	.69	*	*	*	.72	*	*
Observation A: MC/CA/NARR/CRED	.97	*	.24	*	.72	*	*	*	.76	*	*	*
Observation B: MC/CA/NARR/CRED	.50	*	.32	*	.76	*	*	*	.70	*	*	*
Argument: MC/SA/NO/NARR/NOCRE	*	.30	*	.69	*	.76	*	*	*	.78	*	*
Cornell: MC/CA/NARR/CRED	.10	*	.78	*	.44	*	*	*	.46	*	*	*

Table 6 (continued)

Measure: Format	Model 4A Factors			Model 4B Factors			Model 4C Factors					
	(1) Narrative Context	(2) Non- Narrative Context	(3) Credibility Judgment	(4) No Credibility Judgment	(1) Multiple- Choice Common Alternative	(2) Multiple- Choice Specific Alternative	(3) Constructed- Response Justification	(4) Constructed- Response No Justification	(1) Multiple- Choice Narrative	(2) Multiple- Choice Non- Narrative	(3) Constructed- Response Narrative	(4) Constructed Response Non- Narrative
Oblique Solutions												
Observation A: CR/U/NARR/CRED	.33	*	.57	*	*	*	.64	*	*	*	.64	*
Cornell: CR/U/NARR/CRED	-.18	*	.67	*	*	*	.55	*	*	*	.54	*
Reading UFOs: CR/U/NO/NARR/NO/CRED	*	.35	*	.49	*	*	*	.68	*	*	*	.66
Reading Money: CR/U/NO/NARR/NO/CRED	*	.81	*	.50	*	*	*	.73	*	*	*	.74
Reading News: CR/U/NARR/NO/CRED	-.03	*	*	.51	*	*	*	.53	*	*	.46	*
Ennis-Weir: CR/U/NO/NARR/CRED	*	.11	.43	*	*	*	.44	*	*	*	*	.48
Induction: MC/CA/NARR/CRED	.03	*	.47	*	.50	*	*	*	.50	*	*	*
Reading UFOs: MC/SA/NO/NARR/NO/CRED	*	-.06	*	.68	*	.65	*	*	*	.64	*	*
Reading Money: MC/SA/NO/NARR/NO/CRED	*	.01	*	.50	*	.49	*	*	*	.48	*	*
Reading News: MC/SA/NARR/NO/CRED	.10	*	*	.57	*	.55	*	*	.54	*	*	*
Watson-Glaser: MC/SA/NO/NARR/NO/CRED	*	.17	*	.65	*	.67	*	*	*	.68	*	*
Observation A: MC/CA/NARR/CRED	.62	*	.66	*	.72	*	*	*	.72	*	*	*
Observation B: MC/CA/NARR/CRED	.31	*	.55	*	.63	*	*	*	.59	*	*	*
Argument MC/SA/NO/NARR/NO/CRED	*	.17	*	.76	*	.81	*	*	*	.82	*	*
Cornell: MC/CA/NARR/CRED	-.30	*	.72	*	.56	*	*	*	.56	*	*	*
Correlations Among Factors												
(1)	—				—				—			
(2)	.24	—			.81	—			.84	—		
(3)	*	*	—		.98	.83	—		.99	.82	—	
(4)	*	*	.79	—	.63	.82	.77	—	.68	.83	.81	—

* Parameter constrained to be zero.

interval indicates that the oblique models 4B and 4C have reasonably good fit with the data, and that Model 4A is bordering on an excellent fit. Examining Γ_2 values and its confidence interval leads to the same conclusion. However, the lower bounds of the 90% confidence interval is less than .90 for both models 4B and 4C, reducing the confidence in the judgment of good fit. As with the three-factor models, none of the four-factor models fits the data excellently, although the oblique solution to Model 4A is close.

Five-Factor Models

Table 7 presents the factor loadings, and in the case of correlated factors the correlations between format factors for Models 5A, 5B, and 5C.

In the orthogonal solution to Model 5A, all measures load heavily on the Common Critical Thinking factor. For the Narrative factor, the Reading News measure loads negatively in both its constructed-response and multiple-choice formats, although the loading for the multiple-choice measure is close to zero. In the Nonnarrative factor, all measures load positively, although the loading for the Reading Money multiple-choice measure is close to zero. Both the Credibility Judgment and No Credibility Judgment factors are bipolar.

The oblique solution to Model 5A could not be tested with all correlations among factors unconstrained without finding out-of-bounds estimates for some factor loadings. Attempts to solve this problem by placing further constraints on the model were made until the model reported was reached. In this solution, the bipolarity in the Credibility Judgment and No Credibility Judgment factors remains.

In the orthogonal solutions to both Models 5B and 5C, there were some near-zero loadings for most factors. In the oblique solutions, the factors were highly intercorrelated, especially for Model 5B. The oblique solution to Model 5B tended to have smaller loadings on the common Critical Thinking factor, and higher loadings on the other four factors, compared with the orthogonal solution to Model 5B. In both Models 5B and 5C, there were fewer near-zero loadings in the oblique solutions than in the orthogonal solutions.

Table 9 displays the results of the goodness of fit tests. In the cases of Models 5B and 5C, the values of the χ^2 lead to a rejection of the hypotheses of perfect fit between the observed and computed correlation matrices. Examining R^* for these models indicates reasonably good fit with the data. Examining Γ_2 values leads to the same conclusion. However, the lower bounds of the 90% confidence interval for the orthogonal solutions are less than .90, reducing the confidence in the judgment of good fit.

In the case of Model 5A, the value of the χ^2 means that the hypothesis of perfect fit between the observed and computed correlation matrices cannot be rejected. Furthermore, the values and confidence intervals for the R^* and Γ_2 indices support the conclusion of an excellent fit to the data for both the orthogonal and oblique solutions.

Six-Factor Models

Table 8 presents the factor loadings, and in the case of correlated factors the correlations between format factors for Models 6A, 6B, and 6C. Note that oblique solutions for Models 6A and 6C were not possible because the models provided insufficient constraints for EzPATH to reach convergence.

Table 7
Factor Loadings for Five-Factor Models

Measure: Format	Model 5A Factors					Model 5B Factors					Model 5C Factors				
	(1) Common Critical Thinking	(2) Narrative Context	(3) Non- Narrative Context	(4) Credibility Judgment	(5) No Credibility Judgment	(1) Common Critical Thinking	(2) Multiple- Choice Common Alternative	(3) Multiple- Choice Specific Alternative	(4) Constructed- Response Justification	(5) Constructed- Response No Justification	(1) Common Critical Thinking	(2) Multiple- Choice Narrative	(3) Multiple- Choice Non- Narrative	(4) Constructed- Response Narrative	(5) Constructed- Response Non- Narrative
Orthogonal Solutions															
Observation A: CR/U/NARR/CRED	.62	.14	*	-.17	*	.57	*	*	-.02	*	.60	*	*	.04	*
Cornell: CR/U/NARR/CRED	.54	.25	*	.39	*	.51	*	*	.09	*	.55	*	*	-.09	*
Reading UFOs: CR/U/NO/NARR/NO/CRED	.48	*	.50	*	-.21	.51	*	*	*	.86	.52	*	*	*	.29
Reading Money: CR/U/NO/NARR/NO/CRED	.48	*	.58	*	-.00	.60	*	*	*	.27	.58	*	*	*	.82
Reading News: CR/U/NARR/NO/CRED	.69	-.70	*	*	-.18	.49	*	*	*	.16	.51	*	*	.86	*
Ennis-Weir: CR/U/NO/NARR/CRED	.42	*	.21	.23	*	.47	*	*	.88	*	.48	*	*	*	.04
Induction: MC/CA/NARR/CRED	.44	.12	*	.07	*	.45	.15	*	*	*	.45	.15	*	*	*
Reading UFOs: MC/SA/NO/NARR/NO/CRED	.54	*	.18	*	.33	.63	*	.10	*	*	.60	*	.25	*	*
Reading Money: MC/SA/NO/NARR/NO/CRED	.45	*	.03	*	.34	.47	*	.19	*	*	.46	*	.14	*	*
Reading News: MC/SA/NARR/NO/CRED	.54	-.04	*	*	.29	.55	*	.84	*	*	.57	-.05	*	*	*
Watson-Glaser: MC/SA/NO/NARR/NO/CRED	.51	*	.40	*	.27	.65	*	-.06	*	*	.57	*	.56	*	*
Observation A: MC/CA/NARR/CRED	.73	.37	*	-.38	*	.61	.31	*	*	*	.64	.29	*	*	*
Observation B: MC/CA/NARR/CRED	.51	.41	*	-.12	*	.49	.87	*	*	*	.49	.87	*	*	*
Argument: MC/SA/NO/NARR/NO/CRED	.64	*	.39	*	.27	.80	*	-.06	*	*	.74	*	.23	*	*
Cornell: MC/CA/NARR/CRED	.58	.18	*	.51	*	.53	.05	*	*	*	.56	.04	*	*	*

Table 7 (continued)

Measure: Format	Model 5A Factors					Model 5B Factors					Model 5C Factors				
	(1) Common Critical Thinking	(2) Narrative Context	(3) Non- Narrative Context	(4) Credibility Judgment	(5) No Credibility Judgment	(1) Common Critical Thinking	(2) Multiple- Choice Common Alternative	(3) Multiple- Choice Specific Alternative	(4) Constructed- Response Justification	(5) Constructed- Response No Justification	(1) Common Critical Thinking	(2) Multiple- Choice Narrative	(3) Multiple- Choice Non- Narrative	(4) Constructed- Response Narrative	(5) Constructed- Response Non- Narrative
Oblique Solutions															
Observation A: CR/U/NARR/CRED	.63	-.18	*	-.09	*	-.39	*	*	.51	*	.46	*	*	.89	*
Cornell: CR/U/NARR/CRED	.50	-.21	*	.46	*	-.00	*	*	.66	*	.62	*	*	.06	*
Reading UFOs: CR/U/NO/NARR/NO/CRED	.49	*	.46	*	-.15	-.22	*	*	*	.65	.41	*	*	*	.46
Reading Money: CR/U/NO/NARR/NO/CRED	.48	*	.63	*	-.03	-.24	*	*	*	.68	.37	*	*	*	.84
Reading News: CR/U/NARR/NO/CRED	.72	.68	*	*	-.13	-.08	*	*	*	.53	.47	*	*	.14	*
Ennis-Weir: CR/U/NO/NARR/CRED	.40	*	.21	.26	*	.07	*	*	.51	*	.49	*	*	*	.15
Induction: MC/CA/NARR/CRED	.43	-.13	*	.14	*	-.08	.48	*	*	*	.47	.11	*	*	*
Reading UFOs: MC/SA/NO/NARR/NO/CRED	.53	*	.16	*	.40	-.11	*	.64	*	*	.59	*	.25	*	*
Reading Money: MC/SA/NO/NARR/NO/CRED	.44	*	.05	*	.29	-.22	*	.45	*	*	.40	*	.26	*	*
Reading News: MC/SA/NARR/NO/CRED	.53	.01	*	*	.27	-.22	*	.51	*	*	.52	.17	*	*	*
Watson-Glaser: MC/SA/NO/NARR/NO/CRED	.49	*	.39	*	.32	-.05	*	.69	*	*	.51	*	.48	*	*
Observation A: MC/CA/NARR/CRED	.75	-.44	*	-.28	*	-.86	.52	*	*	*	.47	.87	*	*	*
Observation B: MC/CA/NARR/CRED	.50	-.43	*	-.01	*	-.36	.50	*	*	*	.44	.41	*	*	*
Argument MC/SA/NO/NARR/NO/CRED	.62	*	.38	*	.30	-.18	*	.78	*	*	.65	*	.49	*	*
Cornell: MC/CA/NARR/CRED	.53	-.13	*	.57	*	.10	.74	*	*	*	.67	-.02	*	*	*
Correlations Among Factors															
(1)	—					—					—				
(2)	*	—				*	—				*	—			
(3)	*	-.06	—			*	.75	—			*	.30	—		
(4)	*	*	*	—		*	.96	.79	—		*	.47	.19	—	
(5)	*	*	*	.26	—	*	.54	.81	.71	—	*	.30	.66	.28	—

* Parameter constrained to be zero.

Table 8
Factor Loadings for Six-Factor Models

Measure: Format	Model 6A Factors						Model 6B Factors						Model 6C Factors					
	(1) Mul. Ch.	(2) Const. Resp.	(3) Narr.	(4) Non- Narr.	(5) Cred. Judg.	(6) No Cred. Judg.	(1) Narr.	(2) Non- Narr.	(3) Mul. Ch.	(4) Mul. Ch.	(5) Const. Resp.	(6) Const. Resp. No Just.	(1) Cred. Judg.	(2) No Cred. Judg.	(3) Mul. Ch.	(4) Mul. Ch.	(5) Const. Resp.	(6) Const. Resp. Non-Narr.
Orthogonal Solutions																		
Observation A: CR/U/NARR/CRED	.58	-.03	.62	.22	.56	.62	.62	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Cornell: CR/U/NARR/CRED	.30	.07	.56	.68	.56	.62	.56	.50	.33	.16	.03	.71	.65	.60	.52	.49	.37	-.09
Reading UFOs: CR/U/NO/NARR/NO/CRED	.48	.82	.25	.68	.25	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.28
Reading Money: CR/U/NO/NARR/NO/CRED	.64	.14	.30	.68	.30	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Reading News: CR/U/NARR/NO/CRED	.08	.20	.55	.68	.55	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.14
Ennis-Weir: CR/U/NO/NARR/CRED	.36	-.03	.47	.34	.47	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.07
Induction: MC/CA/NARR/CRED	.14	.14	.47	.19	.47	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Reading UFOs: MC/SA/NO/NARR/NO/CRED	.21	.37	.45	.21	.45	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Reading Money: MC/SA/NO/NARR/NO/CRED	.09	.25	.44	.09	.44	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Reading News: MC/SA/NARR/NO/CRED	.21	.19	.35	.21	.35	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Watson-Glaser: MC/SA/NO/NARR/NO/CRED	.39	.30	.42	.39	.42	.62	.50	.50	.33	.16	.03	.71	.62	.52	.78	.78	.78	.78
Observation A: MC/CA/NARR/CRED	.94	.26	.78	-.03	.78	.62	.50	.50	.33	.16	.03	.71	.68	.60	.52	.49	.37	.73
Observation B: MC/CA/NARR/CRED	.40	.48	.56	.08	.56	.62	.50	.50	.33	.16	.03	.71	.54	.60	.52	.49	.37	.27
Argument: MC/SA/NO/NARR/NO/CRED	.44	.44	.37	.37	.37	.62	.50	.50	.33	.16	.03	.71	.70	.77	.77	.77	.77	.08
Cornell: MC/CA/NARR/CRED	.19	.24	.67	.66	.67	.62	.50	.50	.33	.16	.03	.71	.70	.77	.77	.77	.77	-.25

Table 8 (continued)

Measure: Format	Model 6A Factors						Model 6B Factors						Model 6C Factors					
	(1) Mul. Ch.	(2) Const. Resp.	(3) Narr.	(4) Non- Narr.	(5) Cred. Judg.	(6) No Cred. Judg.	(1) Narr.	(2) Non- Narr.	(3) Mul. Ch.	(4) Mul. Ch.	(5) Const. Resp. Just.	(6) Const. Resp. No Just.	(1) Cred. Judg.	(2) No Cred. Judg.	(3) Mul. Ch.	(4) Mul. Ch.	(5) Const. Resp.	(6) Const. Resp. Non-Narr.
Oblique Solutions																		
Observation A: CR/U/NARR/CRED							.39	*	*	*	.56	*						
Cornell: CR/U/NARR/CRED							-.11	*	*	*	.68	*						
Reading UFOs: CR/U/NO/NARR/NO/CRED							*	.38	*	*	*	.55						
Reading Money: CR/U/NO/NARR/NO/CRED							*	.68	*	*	*	.49						
Reading News: CR/U/NARR/NO/CRED							.01	*	*	*	*	.63						
Ennis-Weir: CR/U/NO/NARR/CRED							*	.13	*	*	.46	*						
Induction: MC/CA/NARR/CR							.06	*	.47	*	*	*						UNABLE TO TEST MODEL
Reading UFOs: MC/SA/NO/NARR/NO/CRED							*	.08	*	.67	*	*						
Reading Money: MC/SA/NO/NARR/NO/CRED							*	.10	*	.49	*	*						
Reading News: MC/SA/NARR/NO/CRED							.14	*	*	.58	*	*						
Watson-Glaser: MC/SA/NO/NARR/NO/CRED							*	.31	*	.60	*	*						
Observation A: MC/CA/NARR/CRED							.67	*	.61	*	*	*						
Observation B: MC/CA/NARR/CRED							.35	*	.53	*	*	*						
Argument MC/SA/NO/NARR/NO/CRED							*	.35	*	.71	*	*						
Cornell: MC/CA/NARR/CRED							-.28	*	.78	*	*	*						
Correlations Among Factors																		
(1)							—											
(2)							.37	—										
(3)							*	*	—									
(4)							*	*	.75	—								
(5)							*	*	.92	.78	—							
(6)							*	*	.62	.78	.76	—						

Table 9
Fit Indices for Hypothesized Factor Models

Model	df	χ^2 (prob.)	Steiger-Lind Adjusted RMS Index (R*) (90% confidence interval)	Adjusted Population Gamma Index (Γ_2) (90% confidence interval)
Three-Factor Models				
Orthogonal factors				
3A	75	164.3 (.00)	.080 (.062-.097)	.904 (.862-.941)
3B	75	134.1 (.00)	.065 (.045-.084)	.935 (.896-.968)
3C	75	127.3 (.00)	.062 (.042-.081)	.941 (.902-.973)
Oblique Factors				
3A	74	143.7 (.00)	.073 (.054-.091)	.920 (.879-.955)
3B	74	129.0 (.00)	.063 (.043-.082)	.939 (.899-.972)
3C	74	127.4 (.00)	.064 (.043-.082)	.938 (.898-.971)
Four-Factor Models				
Orthogonal factors				
4A	75	198.3 (.00)	.087 (.069-.104)	.889 (.845-.927)
4B	90	425.9 (.00)	.154 (.140-.168)	.704 (.662-.746)
4C	90	454.9 (.00)	.161 (.147-.176)	.682 (.640-.724)
Oblique Factors				
4A	73	109.5 (.00)	.050 (.024-.070)	.961 (.925-.991)
4B	84	171.2 (.00)	.076 (.058-.093)	.914 (.875-.948)
4C	84	180.5 (.00)	.079 (.062-.096)	.906 (.867-.941)
Five-Factor Models				
Orthogonal Factors				
5A	60	73.1 (.12)	.034 (.000-.060)	.982 (.944-1.000)
5B	75	151.3 (.00)	.072 (.053-.090)	.921 (.880-.956)
5C	75	159.1 (.00)	.077 (.059-.095)	.910 (.868-.946)
Oblique Factors				
5A	58	71.7 (.11)	.034 (.000-.060)	.982 (.943-1.000)
5B	69	109.3 (.00)	.055 (.031-.075)	.953 (.914-.985)
5C	69	117.0 (.00)	.061 (.039-.081)	.942 (.902-.976)
Six-Factor Models				
Orthogonal Factors				
6A	60	152.7 (.00)	.092 (.072-.110)	.874 (.823-.920)
6B	75	225.6 (.00)	.096 (.079-.113)	.865 (.820-.906)
6C	75	235.0 (.00)	.099 (.083-.116)	.856 (.811-.898)
Oblique Factors				
6A	68	99.0 (.01)	UNABLE TO TEST MODEL	.968 (.930-.997)
6B			.046 (.015-.068)	
6C			UNABLE TO TEST MODEL	

In the orthogonal solution to Model 6A, the Multiple-choice and No Credibility Judgment factors were the most clearly defined. The other four factors each contained a mixture of substantial and almost zero loadings. In the orthogonal solution to Model 6B all factors are reasonably clearly defined; the Multiple-choice Common Alternative factor was bipolar. The oblique solution to Model 6B removed the bipolarity on the Multiple-choice Common Alternative factor and introduced it into the Narrative Context factor. Substantial correlations were found among those factors that were allowed to be correlated in the analysis. In the orthogonal solution to Model 6C, the Credibility Judgment and No Credibility Judgment factors were the most clearly defined. Each of the remaining four factors contained at least one near-zero loading, and the Multiple-choice Narrative factor was bipolar.

Table 9 displays the results of the goodness of fit tests. For all the models tested, the value of the χ^2 leads to a rejection of the hypothesis of perfect fit between the observed and computed correlation matrices. Only in the case of the oblique solution to Model 6B did the R^* and Γ_2 values indicate excellent fit, but this judgment must be tempered by the fact that the lower bound of the 90% confidence interval for Γ_2 is less than .95, and the upper bound of the interval on R^* is greater than .05.

Discussion

Twelve models were hypothesized to explain the correlations among 15 measures of critical thinking on the basis of the different formats of the measures. Each of the models was translated into a hypothesized factor structure that was then tested using confirmatory factor analysis. Except for two cases (Models 6A and 6C), each factor structure was tested, first, with the factors constrained to be orthogonal and, second, with at least some of the factors allowed to be correlated. Oblique solutions to Models 6A and 6C could not be found.

Only one model, 5A, was able to pass the χ^2 test of perfect fit between the observed correlations among the measures and the correlations implied by the hypothesized factor structure. The model also showed excellent fit to the data according to both the Steiger-Lind Adjusted RMS Index and the Adjusted Population Gamma Index. According to both these indices it was the best fitting model. The orthogonal and oblique solutions performed the same according to all three indices. The factor loadings also were very similar for both solutions, so the criterion of theoretical simplicity would suggest support of the orthogonal solution.

Model 5A postulated five factors: a Common Critical Thinking factor, and four simple format factors: Narrative Context, Nonnarrative Context, Credibility Judgment Required, and No Credibility Judgment Required. The Common Critical Thinking factor received substantial loadings from all measures, which suggests that all these measures to a large extent test for the same construct. This is an interesting and perhaps happy finding given the diversity of the measures and the conceptual models on which they are based and the fact that they purport to measure aspects of the same construct, critical thinking. Of the remaining four factors two refer to the context (narrative versus nonnarrative) in which the items were set, and two to the nature of the task presented to examinees (credibility judgment required versus no credibility judgment re-

quired). Three of these four factors were bipolar. This presents interpretation problems because conceptually each of the factors was designed as a contrasting pair. Thus, for example, what does it mean to have a negative loading on the Narrative Context factor as opposed to a positive loading on the Nonnarrative Context factor? The bipolarity challenges the concept labels that were placed on the factors and used to construct the Model 5A hypothesis. Thus although the result was a model that fitted the data well, the underlying empirical structure of the model perhaps is not explained well by the conceptual structure that was used to design it.

Among the three-factor and four-factor models, Models 3C (orthogonal solution), and 4A (oblique solution) showed the best fit to the data on the basis of both the Steiger-Lind and Population Gamma indices. We see in Models 3C and 4A only factors also contained in Model 5A. Taken together these results suggest the rejection of the Multiple-choice and Constructed-response factors as viable considerations in attempting to explain the correlations among the set of critical thinking measures studied here. This conclusion is significant, because most of the debate over the format for critical thinking tests has centered around the relative adequacy of multiple-choice and constructed-response approaches. The results of this study suggest that much of this debate is groundless because that particular format difference had little distinguishable effect on performance in the measures studied.

Models 4B, 5B, and 6B (oblique solutions) were exceptions to this finding. These models did not fit the data as well as Model 5A, but showed a reasonably good fit on the basis of the Steiger-Lind and Population Gamma indices. The models employed compound Multiple-choice and Constructed-response factors. Thus an effect on performance for the multiple-choice versus constructed-response format difference appears when other task differences in the multiple-choice and constructed-response formats are represented in the factor structure. Given this finding, and the finding that Model 3A, which contained simple multiple-choice and constructed-response factors, tested less well than Models 3B and 3C, it is reasonable to conclude that the overriding determiners in Models 4B, 5B, and 6B were other than the multiple-choice and constructed-response formats.

One of the major limitations of this study arises from the selection of critical thinking measures examined. The 15 measures originate from only four organizations. The precaution was taken to test the hypothesis that four organization-of-origin factors could account for the data. The orthogonal solution to this model had to be rejected on the basis of all three indices of fit. The oblique solution also had to be rejected, although it showed a reasonably good fit to the data on the basis of the Steiger-Lind index and fitted the data as well as some of the other rejected models. Nevertheless, the measures do represent a narrowness of origin that limits the generalizability of the results. The narrowness was hardly avoidable, however, because few critical thinking tests are available commercially (Norris & Ennis, 1989), and those selected for the study represent the available measures well.

In addition to limited range of origin, the instruments also represent a limited range of content. In particular, there were no subject-specific critical thinking measures studied, that is, measures of critical thinking in, for ex-

ample, science, mathematics, literature, or history. It is impossible to predict how the results would have turned out if such measures had been available and represented in the study. Subject matter content would have suggested other explanatory factors to explore. Little is known about critical thinking testing in specific subjects.

The study was limited to performance by senior high school students, but the range of advertised applicability for some of the measures extended below or above high school grades, or both. Thus the results cannot be generalized automatically to all subjects with whom these measures might be used. Perhaps, for instance, younger students are differently influenced by format effects of the type studied here than older students.

Some precaution is warranted in interpreting the results of confirmatory factor analysis in terms of measurement theory concepts. There is a similarity between the classical test theory decomposition of variable scores into true and error scores, and the factor analytic decomposition into common and unique factors. It is this similarity that has led to the use of confirmatory factor analysis as a technique for checking hypotheses about test equivalence. Thus the definition of parallel measures can be translated into factor analytic language as: "two tests are said by psychometricians to be parallel if they share equal amounts of a common factor, and each also has the same amount of specific variance" (Loehlin, 1987, p. 85). By extension, tests would be tau-equivalent if they share equal amounts of a common factor, but do not have the same amount of specific variance; and tests would be congeneric if they share a common factor, but not to the same degree, and do not have the same amount of specific variance. On this interpretation this study conducted tests of whether groups of measures were congeneric.

There are some theoretical problems with equating classical test theory and common factor theory. Even though the foundations of common factor theory and classical test theory have been laid for almost 90 years, there is no full understanding of the theoretical and empirical relationships between partitioning scores into true and error scores on the one hand, and common and unique factors on the other (McDonald, 1985). The problem arises because it is not theoretically sound to assume that all the unique factor variance in scores is attributable to error. For example, part of the unique factor variance of a test could be attributable to factors that the test shares with other tests that currently are not under consideration. This means that when confirmatory factor analysis is used to check hypotheses of test equivalence, it must be understood that the analysis does not distinguish between variance that is unique to a test in the context of the current analysis but that it shares with other tests not under consideration, and variance that is error and thus uncorrelated with anything else.

Having examined the above limitations, it also must be said that the factor structure of Model 5A makes conceptual and theoretical sense. The existence of a Common Critical Thinking factor should have been expected if the measures were to have theoretical unity. All 15 measures purport to examine some aspects of critical thinking, and it is reasonable to assume that different aspects of critical thinking are highly related to one another. A key word in the previous sentence is *purport*. It must be kept in mind that there was no inde-

pendent test of the construct validity of the measures used in this study. So the finding of a common factor, strictly speaking, demonstrates only that all the measures test *something* in common. The finding does not, in addition, mean that this something is critical thinking. Calling it *critical thinking* is, however, a reasonable hypothesis given two facts: each measure comes with evidence from its developer that it measures critical thinking, and the measures demonstrably test something in common.

The Narrative and Nonnarrative Context factors also make sense. It is a consistent finding that narrative texts are easier to understand than nonnarratives for both adults and children (Bereiter & Scardamalia, 1982; Bock & Brewer, 1985). Narrative is the text type used for initial reading instruction and remains thereafter the type most comprehensible to readers. Clearly initial comprehension of the text with clarity and precision is fundamental to thinking critically about the content.

The Credibility Judgment and No Credibility Judgment factors also make sense as determiners of performance on these measures. Credibility judgment is a sophisticated form of thought requiring knowledge of specific criteria of adjudication and skill in the use and weighing and balancing of these criteria. If an examinee does not know or think to use suitable criteria, then he or she cannot perform well.

This study presents solid evidence to help temper much of the rhetoric surrounding critical thinking testing (McPeck, 1981; Whimbey, 1985). In particular the results are relevant to the dispute over the relative merits of multiple-choice and constructed-response critical thinking measures. In line with Hogan's (1981) conclusion from an examination of this format difference in other than critical thinking tests, the two formats appear to be equivalent or nearly so. In line with Traub and MacRury (1990), any difference that exists is slight. On the other hand, the type of text in which items are cast and whether or not examinees are asked to judge credibility surface as important factors. The latter task is highly significant because facility with it is at the foundation of all critical thinking.

Acknowledgment

Preparation of this manuscript was supported by a grant from the Social Sciences and Humanities Research Council of Canada, grant number 410-89-0052.

References

- Arter, J., & Salmon, J. (1987). *Assessing higher order thinking skills: A consumer's guide*. Portland, OR: Northwest Regional Educational Laboratory.
- Barnett-Foster, D., & Nagy, P. (1995). A comparison of undergraduate test response strategies for multiple-choice and constructed-response questions. *Alberta Journal of Educational Research*, 41, 18-35.
- Bennett, R.E., & Ward, W.C. (Eds.). (1993). *Construction versus choice in cognitive measurement*. Hillsdale, NJ: Erlbaum.
- Bereiter, C., & Scardamalia, M. (1982). From conversation to composition. The role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology* (vol. 2, pp. 1-64). Hillsdale, NJ: Erlbaum.
- Black, M. (1946). *Critical thinking*. Englewood Cliffs, NJ: Prentice-Hall.
- Bock, J.K., & Brewer, W.F. (1985). Discourse structure and mental models. In T.H. Carr (Ed.), *The development of reading skills* (pp. 55-75). San Francisco, CA: Jossey-Bass.
- Dewey, J. (1933). *How we think*. Boston: Heath.
- Ennis, R.H. (1962). A concept of critical thinking. *Harvard Educational Review*, 32, 81-111.

- Ennis, R.H. (1981). Rational thinking and educational practice. In J.F. Soltis (Ed.), *Philosophy of education* (pp. 143-83). Chicago, IL: National Society for the Study of Education.
- Ennis, R.H. (1984). Problems in testing informal logic/critical thinking/reasoning ability. *Informal Logic*, 6(1), 3-9.
- Ennis, R.H., & Millman, J. (1985). *Cornell critical thinking test level X*. Pacific Grove, CA: Midwest.
- Ennis, R.H., Millman, J., & Tomko, T.N. (1985). *Cornell critical thinking tests levels X and Z manual*. Pacific Grove, CA: Midwest.
- Ennis, R.H., & Norris, S.P. (1990). Critical thinking assessment: Status, issues, needs. In S. Legg & J. Algina (Eds.), *Cognitive assessment of language and math outcomes* (pp. 1-42). Norwood, NJ: Ablex.
- Ennis, R.H., & Weir, E. (1985). *The Ennis-Weir critical thinking essay test*. Pacific Grove, CA: Midwest.
- Heim, A.W., & Watts, K.P. (1967). An experiment on multiple-choice versus open-ended answering in a vocabulary test. *British Journal of Educational Psychology*, 37, 339-346.
- Hogan, T.P. (1981). *Relationship between free-response and choice-type tests of achievement: A review of the literature*. Green Bay, WI: University of Wisconsin. (ERIC Document Reproduction Service No. ED 224 811)
- Hullfish, H.G., & Smith, P.G. (1968). *Reflective thinking: The method of education*. New York: Dodd, Mead.
- Kim, J., & Mueller, C.W. (1978). *Introduction to factor analysis*. Beverly Hills, CA: Sage.
- Loehlin, J.C. (1987). *Latent trait models*. Hillsdale, NJ: Erlbaum.
- McDonald, R.P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Erlbaum.
- McPeck, J. (1981). *Critical thinking and education*. New York: St. Martin's.
- Norris, S.P. (1986). *Test on appraising observations (constructed-response version)*. St. John's, NF: Memorial University of Newfoundland.
- Norris, S.P. (1992). A demonstration of the use of verbal reports of thinking in multiple-choice critical thinking test design. *Alberta Journal of Educational Research*, 38, 155-176.
- Norris, S.P., & Ennis, R.H. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Midwest.
- Norris, S.P., & King, R. (1983). *Test on appraising observations (multiple-choice version)*. St. John's, NF: Memorial University of Newfoundland.
- Norris, S.P., & Ryan, J. (1989). *Norris-Ryan argument analysis test*. St. John's, NF: Memorial University of Newfoundland.
- Norris, S.P., & Ryan, J. (1992). *Test of inductive reasoning principles*. St. John's, NF: Memorial University of Newfoundland.
- Perkins, D.N., Jay, E., & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly*, 39, 1-21.
- Phillips, L.M. (1989). *Test of inference ability in reading comprehension (constructed-response format)*. St. John's, NF: Memorial University of Newfoundland.
- Phillips, L.M., & Patterson, C.C. (1987). *Test of inference ability in reading comprehension (multiple-choice format)*. St. John's, NF: Memorial University of Newfoundland.
- Resnick, L.B. (1987). *Education and learning to think*. Washington, DC: National Academy Press.
- Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and education*. London: Routledge.
- Smith, B.O. (1953). The improvement of critical thinking. *Progressive Education*, 30(5), 129-134.
- Steiger, J.H. (1989). *Causal modeling*. Evanston, IL: SYSTAT.
- Traub, R.E. (1993). On the equivalence of the traits assessed by multiple-choice and constructed-response tests. In R.E. Bennett & W.C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 29-44). Hillsdale, NJ: Erlbaum.
- Traub, R.E., & Fisher, C.W. (1977). On the equivalence of constructed-response and multiple-choice tests. *Applied Psychological Measurement*, 1, 355-369.
- Traub, R.E., & MacRury, K. (1990). Antwort-auswahl vs. freie-antwort-aufgaben bei Lernervolks-tests [Multiple-choice vs. free-response in the testing of scholastic achievement]. In K. Ingenkamp & R.S. Jäger (Eds.), *Tests und trends 8: Jahrbuch der pädagogischen diagnostik* (pp. 128-159). Weinheim, Germany: Beltz Verlag.
- Watson, G., & Glaser, E.M. (1980). *Watson-Glaser critical thinking appraisal*. Cleveland, OH: Psychological Corporation.
- Whimbey, A. (1985). You don't need a special "reasoning" test to implement and evaluate reasoning training. *Educational Leadership*, 43(2), 37-39.

José L. da Costa
University of Alberta

Teacher Collaboration: A Comparison of Four Strategies

Data for eight constructs were collected from 26 teachers and their 401 students in two public school districts in the British Columbia Lower Mainland. The question was: What teacher and pupil characteristics differentiate the four teacher collaboration strategies consisting of collaborative consultation, collaborative consultation with team teaching, collaborative consultation with no classroom observation, and collegial consultation with no classroom observation? Data collected for the eight constructs were analyzed using multivariate analysis of variance (MANOVA) followed by post-hoc discriminant analysis (DA). Results of the MANOVA suggest that differences did exist among the four collaboration groups. The post-hoc DA indicates that the groups differ on a dimension anchored by personal teaching efficacy and pupil achievement at one end and pupil attitude at the other. It is concluded that classroom observation is an essential element of the teacher collaboration process if increased pupil achievement and increased teacher efficacy are desired outcomes.

Des données de huit modèles théoriques ont été recueillies de 26 enseignant(e)s et leurs 401 élèves de deux districts scolaires au sud de la Colombie-Britannique. La question posée était qu'elles caractéristiques chez les enseignant(e)s et les élèves différencient les quatre stratégies de collaboration professionnelle suivantes: la consultation collaborative, la consultation collaborative dans un contexte d'enseignement en équipe, la consultation collaborative sans observation directe en salle de classe, et la consultation collégiale n'ayant aucune observation en salle de classe. Les données recueillies pour les huit modèles théoriques ont été analysées selon l'analyse de variantes multivariées (AVAMUL) suivie d'une analyse destinatrice post-hoc (AD). Les résultats de cette première analyse (AVAMUL) suggèrent qu'il existe en effet une différence entre les quatre différents groupes collaboratifs. La deuxième analyse post-hoc (AD) révèle que les groupes diffèrent sur une dimension ancrée d'une part sur l'efficacité d'enseignement personnel et le succès académique des élèves et de l'autre part, sur l'attitude des élèves. Il en a été déterminé que l'observation en salle de classe est un élément essentiel du processus de la collaboration professionnelle si le succès amélioré des élèves et l'efficacité accrue chez les enseignant(e)s sont les résultats désirés.

Over the past two decades educational researchers have conducted many studies investigating the effects on teachers and pupils of various teacher development approaches that emphasize teacher growth (Donovan, Sousa, & Walberg, 1987; Showers, 1985; Smith, 1989; Stallings, 1985). Many of these once-promising teacher development approaches have fallen into disfavor with both the research community and the teachers themselves (Slavin, 1986; Smith & Acheson, 1991; Stallings & Krasavage, 1986). The failure of these teacher development approaches has been attributed to various factors, one of which is lack of teacher commitment to the process due to conflicts between teachers' own norms and values and those imposed externally by the approach taken (Hargreaves & Fullan, 1992). Grimmatt, Rostad, and Ford (1992) echo this sentiment:

José da Costa is an assistant professor in the Department of Educational Policy Studies.

Externally mandated change typically has a cataclysmic effect on teachers' morale, resulting in a strong sense of dependency. Teachers often feel overwhelmed by the new expectations when their actions are continually shaped by the directives of others. There is an accompanying sense of helplessness and powerlessness when heightened expectations appear to be beyond reach. (pp. 185-186)

Consequently, of interest in the study reported in this article was a comparison of teacher consultation approaches the goal of which was to permit teachers to make sense of their classroom behaviors through their own values and norms. This interest, however, was not limited to the effects of the process of consultation on teachers but also examined the effects of the process of teacher consultation on students.

Garman (1986) contended that "ultimately the reason teachers and clinical supervisors work together is in order to enhance practice and to make education better for students" (p. 19). Similarly, Fullan and Hargreaves (1991) believed that improving teachers' pedagogic skills and school working conditions in turn would lead to student improvement. In other words, when teachers collaborate, their students should experience positive change in terms of achievement, attitude, and behavior. Presently, however, these links are only conjectures; empirical evidence is lacking (Acheson & Gall, 1992; Greene, 1987; Wildman & Niles, 1987). At a practical level the purpose of this study was to provide some initial direction to educators in developing teacher collaboration programs that benefit teachers as well as their pupils. Given the myriad possibilities in which teachers can work with other teachers, this study sought to collect empirical evidence to address the question: What teacher and pupil characteristics differentiate the four teacher collaboration strategies consisting of collaborative consultation (CC), collaborative consultation with team teaching (CCTT), collaborative consultation with no classroom observation (CCNO), and collegial consultation with no classroom observation (CoNO)? These four groups were chosen for the simple and practical reason of accessibility. Descriptions of the characteristics defining each group are found in Appendix A.

Review of the Literature

The literature suggests that many variables may be linked to teacher collaboration and its effects on teachers and pupils. This brief review explores those teacher and pupil variables that appear to be related to teacher collaboration.

Teacher growth through the use of teacher collaboration techniques is predicated on the establishment of teacher trust for the teaching partner (Acheson & Gall, 1992; Cogan, 1973; Goldhammer, Anderson, & Krajewski, 1980; Grimmett & Erickson, 1988; Hargreaves & Dawe, 1990; Lovell & Wiles, 1983; Sergiovanni & Starratt, 1993). This approach leads to the expectation that teacher trust for the teaching partner is influenced by the supervisory mode preferred by the teaching partner. Using Glickman's (1990) terms, the teaching partner's preferred mode of supervisory interaction should not be directive, but rather nondirective or collaborative if the intent is to promote trust for the teaching partner.

A potentially positive outcome of teacher collaboration when the above conditions are met is an increase in general and personal teaching efficacy

(Ashton, Webb, & Doda, 1983; Cavers, 1988; Denham & Michael, 1981). Cruickshank and Applegate's (1981) conclusions suggest that increased levels of self-esteem and a positive outlook on teaching leads teachers to become more concerned with self-improvement. Presumably, as a consequence of becoming more concerned with self-improvement teachers change their classroom behavior. Increased general and personal teaching efficacy can be thought to give the teacher confidence that improvement can be achieved by trying new approaches to solving classroom problems (Bandura, 1978).

Many authors (Acheson & Gall, 1992; Fullan & Hargreaves, 1991; Little, 1987; Lovell & Wiles, 1983; Sergiovanni & Starratt, 1993) agree that if teachers can be encouraged to think critically about what they do in the classroom and then act on those thoughts, there should be an accompanying change in pupils' achievement, attitudes, and behavior. However, these linkages are theoretical and have not as yet been demonstrated empirically.

This review suggests that at least eight teacher and pupil constructs are of importance for the present investigation of how teacher collaboration affects teachers and pupils. These constructs include: (a) teacher trust for the teaching partner, (b) the teaching partner's supervisory beliefs, (c) general teacher efficacy, (d) personal teacher efficacy, (e) teacher classroom behavior, (f) pupil achievement, (g) pupil attitude, and (h) pupil behavior.

Method

The Participants

The study employed a pre- postmeasurement survey design with no researcher intervention. The sample consisted of 26 self-selected classes (26 teachers and 401 pupil volunteers) at the elementary level from two public school districts in the British Columbia Lower Mainland. All of these teachers had previously participated in the British Columbia Teachers' Federation Program for Quality Teaching, which provided instruction and practice in the use of peer supervision techniques for the purpose of professional development (Smith & Acheson, 1991).

Four consultation strategies—arrived at inductively—describe how the teacher volunteers worked with their teaching partners: (a) collaborative consultation between teacher dyads teaching in separate classrooms, (b) collaborative consultation between team teacher dyads teaching in one double sized classroom, (c) collaborative consultation without direct observation with dyads teaching in separate classrooms, and (d) collegial consultation without direct observation (descriptions of the nature of each group are found in the Appendix). To determine into which of the four groups teachers fitted best each was asked the following questions to determine the nature of the professional relationship between the teacher and teaching partner: (a) Will you collaborate with a teaching partner? (b) In what capacity is your teaching partner employed in the school district? (c) Will your teaching partner be observing your classroom teaching to collect data for conferencing purposes? and (d) Do you share one open classroom with your teaching partner?

The Measures

Data were collected for the eight constructs identified in the literature as relevant. The scales used to obtain measures of each of the eight variables

deemed important by the literature were, respectively, as follows: (a) Wheelless and Grotz's (1977) Individualized Trust Scale (ITS), (b) Glickman and Tamashiro's (1981) Supervisory Beliefs Inventory (SBI), (c) Gibson and Dembo's (1984) Teacher Efficacy Scale—general dimension (TESg), (d) Gibson and Dembo's (1984) Teacher Efficacy Scale—personal dimension (TESp), (e) Eash and Waxman's (1983) Our Class and Its Work (OCIW), (f) pupil achievement grades from district report cards (ACH), (g) Randhawa and Van Hesteren's (1983) School Attitude Scales for Children (SASC), and (h) pupil behavior grades (BEH) from district report cards (see Appendix for details regarding each of these measures).

Data collection from took place at two points during the school year. Premeasure data were collected one month after the beginning of the school year, and the postmeasure data were collected one month before the end of the school year. Identical scales were used for both sets of measurements.

Data Analyses

Because the sample consisted of self-selected volunteers, it was decided that the analyses of the data should be carried out in three phases: (a) the first phase compared the sample with other samples and populations on the basis of five variables, (b) the second phase compared the four groups with one another on the basis of the premeasures data, and (c) the third phase compared the four groups on the basis of the postmeasures data. In order to prevent inflation of statistical power, all analyses used the class as the unit of analysis. This conservative approach was chosen to reduce the possibility of making a type I error in the statistical analyses.

First, descriptive statistics for each variable were calculated to screen the data and gain a preliminary understanding of the data set. Second, group differences on the premeasures were sought. A multivariate analysis of variance (MANOVA) was conducted on all the premeasure variables simultaneously to determine if the four collaboration groups differed near the beginning of the school year. Third, group differences on the postmeasures were sought using MANOVA. This was followed by post-hoc discriminant analysis (DA) to determine how the collaboration groups differed.

Limitations

The limitations of this study fall into two basic areas: (a) the use of a volunteer sample, and (b) the method of data analysis. Each of these is examined in turn.

This study drew data from a sufficiently large sample that, had it been randomly selected, the generalizability of the results to the population of teachers working in the suburban area surrounding Vancouver would not have been problematic. However, because the sample consisted entirely of teacher and pupil volunteers the results are not truly generalizable beyond the population of teachers and pupils who would volunteer for such a study. In other words, it is recognized that the variables associated with teacher collaboration are themselves affected by the degree to which the participants wish to participate. This fact is less problematic than it initially might sound. In discussing how to best implement professional development programs utilizing teacher collaboration, Hargreaves and Fullan (1992) and Greene (1992) both place a great deal of emphasis on the desirability of having teachers choose

whether to participate and, if so, how to participate in such programs. Presumably, the types of data gathered from the volunteer sample in this study would more closely resemble those cases in which teachers decide whether to participate than those cases in which teachers are told that they will collaborate with colleagues.

With respect to the limitations of the method of analysis, Patton (1990) stated "*there are no perfect research designs*" (original emphasis, p. 162). The method employed here is no exception. Analyses making use of MANOVA are more difficult to interpret than corresponding univariate analyses. However, the disadvantage of this limitation is greatly outweighed by the advantage of being able to simultaneously make sense of the differences among several groups using several constructs by taking into account the covariation of the variables of interest (Tabachnick & Fidell, 1989).

Findings

From the preceding analyses three categories of findings emerge, which are best classified as relating to: (a) the sample, (b) the premeasures, and (c) the postmeasures. Each is addressed below.

The Sample

On the basis of teacher experience, female-to-male ratio, teacher efficacy scores (TESg and TEsP), and OCIW scores of teachers and their pupils appear to be reasonably representative of other teachers and pupils in other British Columbia school districts. Using a chi-square goodness of fit test, the teaching experience of the sample teachers was found not to be significantly different ($\alpha = .05$) from the teaching experience of the elementary teachers in the two districts from which the sample was drawn. Similarly, the female-to-male teacher ratio of 2:1 did not differ significantly ($\alpha = .05$) from the British Columbia provincial female-to-male elementary teacher ratio. In the vast majority of school districts, female elementary teachers greatly outnumber male teachers. Comparing the TESg, TEsP, and OCIW mean scores obtained from the present sample with comparable mean scores from previous studies, this sample was seen to be similar to the samples (see Table 1) by Anderson, Greene, and Loewen (1987), Cavers (1988), Tracz and Gibson (1986), Grimmett and Crehan (E.P. Crehan, personal communication, April 22, 1992), and Housego (1992).

The Premeasure Analysis

A MANOVA was conducted to determine if any differences existed among the groups near the beginning of the school year. This MANOVA compared the four collaboration groups simultaneously on the basis of the eight constructs of interest.

After taking all the premeasure variables together to account for the covariation among the variables, the MANOVA revealed no significant differences among the four collaboration groups at the .05 level (Wilks lambda = .331, $F = .853$, $df = 24/44$, $p = .656$). Having established that the groups were probably similar at the beginning of the school year, the decision was made that the postmeasure group data did not need to be adjusted (i.e., through MANCOVA) to compensate for any initial differences among the groups.

Table 1
Comparisons of TES and OCIW Means Between the Present Study and Previous Studies

Study	Teacher Efficacy Personal		Teacher Efficacy General		OCIW Total	
	n	mean	mean	n	mean	
Anderson et al. (1987)	24	4.77	3.56	584	2.70	
Cavers (1988)	339	4.64	3.44			
Tracz & Gibson (1986)	14	4.86	3.65			
Grimmett & Crehan (n.d.)	PRE	93	4.29	3.22		
	POST	78	4.69	3.20		
Housego (1992)	BASE	177	3.97	3.76		
	T1	129	4.38	3.56		
	T2	121	4.50	3.41		
	T3	123	4.47	3.47		
da Costa (1993)	PRE	26	4.56	3.75	401	2.68
	POST	26	4.53	3.62	401	2.68

The Postmeasure Analysis

The results of a MANOVA suggest that differences do exist at the .05 alpha level among the collaboration groups after taking all the postmeasure variables together to account for the covariation among the variables (Wilks lambda = .139, $F = 1.79$, $df = 24/44$, $p = .046$). To find the nature of these differences a post-hoc DA was conducted.

Using the eight postmeasure variables as potential predictors, DA was carried out. At the .05 alpha level, only the first discriminant function was retained ($\chi^2(24) = 37.51$, $p = .039$). Using structure coefficients, with loadings of |.25| or greater, a dimension with pupil attitude (−.44) at one end and personal teaching efficacy (.27) along with pupil achievement (.39) at the other end is described. As can be seen in Figure 1, plotting the discriminant means (CC mean = 1.89, CCTT mean = 0.81, CCNO mean = −1.59, and CoNO mean = 0.00), these findings were obtained: (a) CC group differed most from the other three groups; (b) CCTT group and the CoNO groups were most similar to each other, but they differed from both the CC and the CCNO groups; and (c) the CCNO group differed from all other groups.

The CC group seemed to stand out when compared with other collaboration groups in the present study. Teachers in the CC group exhibited more personal teaching efficacy, and pupils in this group had higher achievement than the teachers and pupils, respectively, in the other collaboration groups. However, pupils in the CC group also tended to have more negative attitudes toward these attitudinal objects addressed in the SASC: themselves, peers, their teacher, school, learning in general, language arts, social studies, and math. Also distinctive in comparison with the other collaboration groups was the CCNO group. Teachers in this group tended to have lower personal teaching

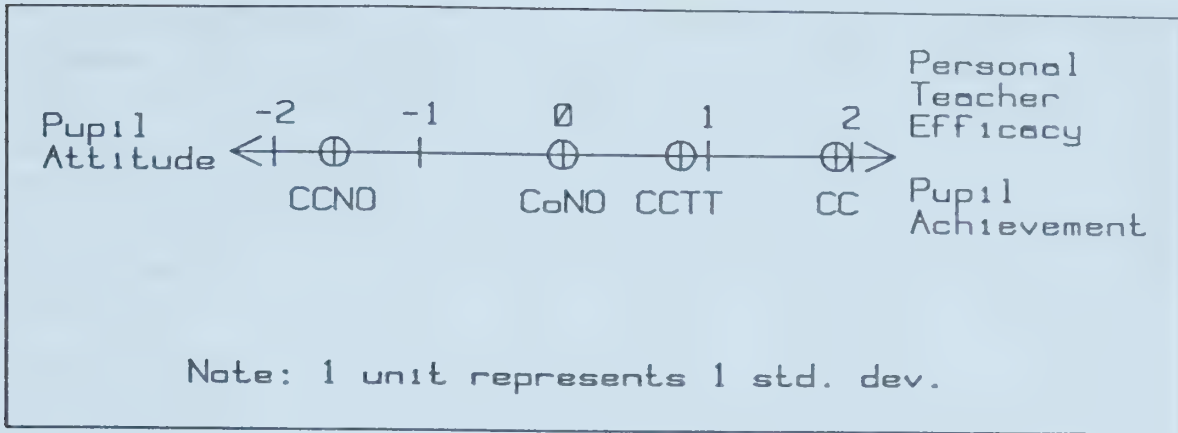


Figure 1. The single bipolar dimension on which the four collaboration groups discriminated.

efficacy and lower pupil achievement, but the same pupils generally had more positive attitudes toward the attitudinal objects addressed by the SASC.

Discussion

It is noteworthy that the post-hoc DA resulted in at least one discriminant function describing a bipolar underlying continuum composed of personal teaching efficacy and pupil achievement at one end of the scale and pupil attitudes at the opposite end. The expectation as a result of the literature reviewed (Acheson & Gall, 1992; Little, 1987) would have been to observe pupil attitudes and pupil achievement anchor at one end of the continuum rather than at opposite ends of the same continuum as found here. This suggests that high achieving pupils generally have more negative attitudes toward attitudinal objects associated with school—at least toward those measured by the SASC.

A second point regarding this continuum involves the relationship between personal teaching efficacy and pupil achievement. Several studies suggest that a reciprocal relationship exists between teaching efficacy and pupil achievement (Armor et al., 1976; Ashton, 1985; Berman, McLaughlin, Bass, Pauly, & Zellman, 1977). Cavers (1988), however, found no evidence of such a relationship. Ashton and Webb (1986) concluded that their findings “strongly support the hypothesis that teachers’ sense of efficacy is related to pupil achievement” (p. 139). The present study suggests that (a) a positive relationship does exist between personal teaching efficacy and pupil achievement as suggested by Anderson et al. (1987), and (b) a negative relationship exists between personal teaching efficacy and pupil attitudes toward school-related attitudinal objects. Furthermore, Canon and Simpson’s (1985) and Randhawa and Van Hesteren’s (1982) conclusion that attitude and achievement are not significantly related is also supported by the present study.

A possible explanation for the grouped data relationships among personal teaching efficacy, pupil achievement, and pupil attitudes may be found by first recalling the definition of personal teaching efficacy and then examining the nine items related to this efficacy dimension on Gibson and Dembo’s (1984) TES. Personal teaching efficacy is described as the belief that the individual teacher has the skills and abilities to bring about pupil learning. This definition is reflected in the personal teaching efficacy items of the TES by the inclusion of

the phrase indicating how the teacher would react to various difficulties. Examples of phrases indicating how the teacher would react to pupil problems include: item 1, "I exerted a little extra effort"; item 5, "I am usually able"; item 6, "I found better ways"; item 13, "I know some techniques." Teachers with high personal teaching efficacy may be more apt than teachers with lower personal teaching efficacy to try various techniques or strategies to increase individual pupil achievement. These techniques or strategies may succeed in increasing pupil achievement; however, pupils may no longer feel they have ownership of the learning process because the teacher has control of it. It may be speculated that this loss of control of the learning process on the part of the pupil may result in more negative pupil attitudes toward school-related attitudinal objects (B.E.J. Housego, personal communication, March 18, 1993).

The general trend of positioning the various group means relative to one another in the discriminant space comprising personal teaching efficacy and pupil achievement at one extreme and pupil attitudes at the other (as seen in Figure 1) was for the most part expected on the basis of what had been indicated in the literature. The expectation was that if groups were going to cluster they would do so according to the following: (a) the collaborative consultation groups would separate from the collegial consultation group, and (b) the groups utilizing direct classroom observation would separate from those groups not utilizing direct classroom observation. As seen in Figure 1 the trend is for the groups using direct classroom observation to cluster as expected (i.e., the collaborative consultation group and the collaborative consultation in a team teaching situation group are positioned side by side in discriminant space). Teachers basing their collaboration conferences on classroom observational data gathered by a peer were better able to effect changes to improve their pupils' levels of achievement. An offshoot of this seems to be increased personal teaching efficacy for the individual teacher. Also as expected, the groups not making use of classroom observation on which to base their follow-up conferences tended to fall on the continuum away from the end anchored by high levels of personal teaching efficacy and pupil achievement.

The collaborative consultation with no observation fell so much farther to the left of the collegial consultation group on this continuum, indicating even lower levels of personal teaching efficacy and pupil achievement. This was surprising because those teachers involved in collegial consultation were not conferring with a peer, but were working in nonreciprocal arrangements with "experts" (i.e., administrators, master teachers) who were in a position to tell the teachers what needed to be done in order to address problems in the teachers' classrooms.

Conclusions

Given the exploratory nature of the present study, the conclusions reached and the recommendations made need to be viewed cautiously. This study is useful from a practical perspective because it suggests some ways for organizing teacher development through collaboration programs and from a theoretical perspective because it provides a first step in obtaining empirical evidence to show how teachers working with other teachers are affected and how their pupils are affected.

The Importance of Classroom Observation

Classroom observation on the part of the teaching partner is an essential element of collaborative consultation if the aim is to increase personal teaching efficacy and pupil achievement. This is what can be expected when collaborative consultation is conducted using Garman's (1986) model of reflection on action which is based on the contextualization by the teacher of "stable" data in light of educational theory through interpretation, explanation, and evaluation.

Collaboration not Based on Classroom Observation

As noted above, conferences based on the data collected from direct classroom observation provide better results when the desire of the teacher is to improve pupil achievement and increase personal teaching efficacy. However, this study suggests that situations that preclude the use of data collection from direct classroom observation are better handled using a nonreciprocal consultation model between a teacher and a pedagogical expert (i.e., administrator, master teacher). Collaborative consultation without direct classroom observation on the part of the teaching partner was the least effective method of the four studied in promoting pupil achievement and personal teacher efficacy.

Relationship Among Personal Teacher Efficacy, Pupil Achievement, and Pupil Attitude

The positive relationship found between personal teacher efficacy and pupil achievement, the negative relationship found between personal teacher efficacy and pupil attitudes, and the negative relationship found between pupil attitudes and achievement seem to contradict some intuitive notions of how these constructs are related. Given this finding and the work of Anderson et al. (1987) and of Ashton and Webb (1986), the conclusion that increases in personal teaching efficacy correspond to increases in pupil achievement seems reasonable. Adding to this the work of Canon and Simpson (1985) and Randhawa and Van Hesteren (1983), who speculated that achievement and attitude were not directly related, it is reasonable to conclude that achievement and attitude are not similarly related to personal teaching efficacy. Rather, increases in personal teaching efficacy correspond to changes in pupil attitude. Furthermore, the mechanism that relates personal teaching efficacy to pupil achievement and attitude is not at this time clearly evident.

Recommendations

In addition for a call to replicate this study, three recommendations follow from the three conclusions drawn above. The first two address practice, the third theory. The recommendations for practice are made under these assumptions (a) teacher participation in a teacher collaboration program is not mandated by district personnel, and (b) teachers select the pedagogic areas in which they wish to change or improve.

Recommendation 1. Teacher collaboration programs in which classroom observation by the teaching partner is an integral part of the design should be encouraged because they seem to be more positively associated with increased personal teacher efficacy and higher pupil achievement than nonobservation collaboration approaches. Typically, teaching partners must be provided re-

lease time from their own classrooms to observe the teacher and his or her class so that the teacher can be provided with what Garman (1984) calls stable data.

Recommendation 2. In situations where teacher collaboration programs are implemented that do not include classroom observation (i.e., the data discussed during conferencing are recalled by the teacher) by the teaching partner as an integral part of the design, teachers should discuss their teaching practices with an expert (i.e., administrator, master teacher) in a nonreciprocal relationship. This is not to say that implementing teacher collaboration programs that do not utilize classroom observation for data collection are recommended, but it is acknowledged that in many circumstances the resources required for this type of program are simply not available.

Recommendation 3. To help understand the mechanism by which personal teaching efficacy, pupil achievement, and pupil attitude are related, a factor analysis of the SASC pupil attitude data should be conducted to determine the nature of the dimensions that underlie the attitude construct. Intuitively, the expectation is that when pupil attitudes change, according to any one of the SASC subscales used in the present study, pupil favorableness toward or against other psychological objects of interest may also change.

Acknowledgements

The feedback provided by Dr. E.A. Holdaway and the reviewers of the draft of this manuscript is gratefully acknowledged.

References

- Acheson, K.A., & Gall, M.D. (1992). *Techniques in the clinical supervision of teachers: Preservice and inservice applications* (3rd ed.). New York: Longman.
- Anderson, R., Greene, M., & Loewen, P. (1987). *Relationships among teachers' and students' thinking skills, sense of efficacy and student achievement*. Unpublished manuscript, University of Lethbridge.
- Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascall, A., Pauly, E., & Zellman, G. (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools*. Santa Monica, CA: Rand.
- Ashton, P.J. (1985). Motivation and the teacher's sense of efficacy. In C. Ames & R. Ames (Eds.), *Research on motivation in education, Volume 2: The classroom milieu* (pp. 141-169). Toronto, ON: Academic Press.
- Ashton, P.J., & Webb, R.B. (1986). *Making a difference*. New York: Longman.
- Ashton, P.J., Webb, R., & Doda, N. (1983). *A study of a teacher's sense of efficacy, final report. Executive summary*. Gainesville, FL: Florida University.
- Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist*, 33, 344-358.
- Berman, P., McLaughlin, M., Bass, G., Pauly, E., & Zellman, G. (1977). *Federal programs supporting educational change. Vol. VII: Factors affecting implementation and continuation*. Santa Monica, CA: Rand.
- Canon, R.K., & Simpson, R.D. (1985). Relationships among attitude, motivation, and achievement of ability grouped, seventh-grade, life science students. *Science Education*, 69(2), 121-138.
- Cavers, L. (1988). *Teacher efficacy: Its relationship to school level organizational conditions and teacher demographic characteristics*. Unpublished doctoral dissertation, University of British Columbia.
- Cogan, M.L. (1973). *Clinical supervision*. Boston, MA: Houghton Mifflin.
- Cruickshank, D.R., & Applegate, J.H. (1981). Reflective teaching as a strategy for teacher growth. *Educational Leadership*, 38(7), 553-554.
- Denham, C., & Michael, J. (1981). Teacher sense of efficacy: A definition of the construct and a model for future research. *Educational Research Quarterly*, 6(1), 39-63.
- Donovan, J.F., Sousa, D.A., & Walberg, H.J. (1987). The impact of staff development on implementation and student achievement. *Journal of Educational Research*, 80, 348-351.
- Eash, M.J., & Waxman, H.C. (1983). Utilizing students' perceptions and context variables to analyze effect teaching: A process-product investigation. *Journal of Educational Research*, 76(6), 321-324.

- Fullan, M., & Hargreaves, A. (1991). *What's worth fighting for? Working together for your school*. Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED 342 128)
- Garman, N.B. (1984, April). *Stable data and clinical supervision*. Paper presented at the annual conference of the Association of Teacher Educators, New Orleans.
- Garman, N.B. (1986). Reflection, the heart of clinical supervision: A modern rationale for professional practice. *Journal of Curriculum and Supervision*, 2(1), 1-24.
- Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, 76(4), 565-582.
- Glickman, C.D. (1990). *Supervision of instruction: A developmental approach* (2nd ed.). New York: Allyn and Bacon.
- Glickman, C.D., & Tamashiro, R.T. (1981). The supervisory beliefs inventory. In C.D. Glickman (Ed.), *Developmental supervision: Alternative practices for helping teachers to improve instruction* (pp. 88-91). Alexandria, VA: Association for Supervision and Curriculum Development.
- Goldhammer, R., Anderson, R.H., & Krajewski, R.J. (1980). *Special methods for the supervision of teachers* (2nd ed.). New York: Holt.
- Greene, M. (1987). *Landscapes of learning*. New York: Teachers College Press.
- Greene, M. (1992). Teacher supervision as professional development: Does it work? *Journal of Curriculum and Supervision*, 7(2), 131-148.
- Grimmett, P.P., & Crehan, E.P. (n.d.). *Collaborative consultation project*. Unpublished manuscript, University of British Columbia.
- Grimmett, P.P., & Erickson, G.L. (1988). *Reflection in teacher education*. Vancouver, BC: University of British Columbia, Pacific Educational Press.
- Grimmett, P.P., Rostad, O.P., & Ford, B. (1992). The transformation of supervision. In C. Glickman (Ed.), *Supervision in transition, 1992 yearbook of the Association for Supervision and Curriculum Development* (pp. 185-202). Alexandria, VA: Association for Supervision and Curriculum Development.
- Hargreaves, A., & Dawe, R. (1990). Paths of professional development: Contrived collegiality, collaborative culture, and the case of peer coaching. *Teaching and Teacher Education*, 6(3), 227-241.
- Hargreaves, A., & Fullan, M. (1992). Shortcomings of traditional approaches to teacher development. *Keys to Teacher Development*, 1(1), 3-4.
- Housego, B.E.J. (1992). Monitoring student teachers' feelings of preparedness to teach, personal teaching efficacy, and teaching efficacy in a new secondary teacher education program. *Alberta Journal of Educational Research*, 38, 49-64.
- Little, J. (1987). Teachers as colleagues. In V.R. Koehler (Ed.), *Educators' handbook: A research perspective* (pp. 491-518). New York: Longman.
- Lovell, J.T., & Wiles, K. (1983). *Supervision for better schools* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Patton, M.Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Randhawa, B., & Van Hesteren, F. (1982, June). *The development and validation of a hybrid semantic differential measure of school related attitudes in young children*. Paper presented at the annual meeting of the Canadian Psychological Association, Montreal. (ERIC Document Reproduction Service No. ED 223 691)
- Randhawa, B., Van Hesteren, F. (1983). *School-related attitudes and their redundancy with academic achievement*. Paper presented at the Learned Societies Conference of the Canadian Society for the Study of Education, Vancouver. (ERIC Document Reproduction Service No. ED 235 227)
- Sergiovanni, T.J., & Starratt, R.J. (1993). *Supervision: Human perspectives* (5th ed.). New York: McGraw Hill.
- Showers, B. (1985). Teachers coaching teachers. *Educational Leadership*, 42(7), 43-48.
- Slavin, R.E. (1986). The Napa evaluation of Madeline Hunter's ITIP: Lessons learned. *Elementary School Journal*, 87(2), 165-171.
- Smith, N.S. (1989). *Peer consultation as staff development. A study in BC Schools*. Unpublished doctoral dissertation, University of Oregon.
- Smith, N.S., & Acheson, K.A. (1991). Peer consultation: An analysis of several types of programs. *Oregon School Study Council*, 34(6), 1-29.
- Stallings, J. (1985). A study of implementation of Madeline Hunter's model and its effects on students. *Journal of Educational Research*, 78(6), 325-337.

- Stallings, J., & Krasavage, E. (1986). The effects of instruction based on the Madeline Hunter model on student's achievement: Findings from a follow-through project. *Elementary School Journal*, 87(2), 571-588.
- Tabachnick, B.G., & Fidell, L.S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper and Row.
- Tracz, S., & Gibson, M. (1986, November). *Effects of efficacy on academic achievement*. Paper presented at the California Educational Research Association Annual Meeting, Marina del Rey. (Eric Document Reproduction Service No. ED 281 853)
- Wheeless, L.R., & Grotz, J. (1977). The measurement of trust and its relationship to self disclosure. *Human Communication Research*, 3(3), 250-257.
- Wildman, T.M., & Niles, J.A. (1987). Reflective teachers: Tensions between abstractions and realities. *Journal of Teacher Education*, 38(4), 25-31.

Appendix A

Description of Terms

Clinical supervision. This refers to the partnership described by Cogan (1973) and Goldhammer et al. (1980) between supervisor and teacher that uses classroom data as the basis for subsequent analyses the purpose of which is to improve the teacher's classroom practices in ways that make sense to the teacher. This process is intended to be for formative purposes only. The supervisor's job is not to identify what is "right" or "wrong" with the teacher's teaching but to help the teacher identify appropriate goals for improvement. The teacher's role is to decide the focus of the clinical supervision process and the direction in which it will proceed.

Collaborative consultation. This also refers to a process intended to facilitate teacher development using the principles of Cogan's (1973) clinical supervision. Underpinning this process, however, is a nonhierarchical relationship between a teacher and a teaching partner characterized by mutual trust and respect that is presumed to provide a supportive environment in which the teacher can evaluate previous teaching strategies as well as implement and evaluate new strategies. It is a relationship of equals in which both partners wish to engage (E.P. Crehan, personal communication, February 15, 1990).

Collaborative consultation without direct observation. This is similar to collaborative consultation as described above except that the classroom observation phase, as described by Cogan (1973), used to collect "objective" data is not implemented; instead, data for conferencing are obtained from the teacher's recollection of past events.

Collegial consultation. This refers to a process intended to facilitate teacher development using the principles of Cogan's (1973) clinical supervision. This is a professional relationship between a teacher and another individual in a school or school district (e.g., principal, master teacher, curriculum specialist, etc.). The relationship is not reciprocal: the two people forming the dyad do not exchange roles.

Collegial consultation without direct observation. This similar to collegial consultation as defined above except that the classroom observation phase, as described by Cogan (1973), used to collect "objective" data is not implemented; instead, data for conferencing are obtained from the teacher's recollection of past events.

Individualized Trust Scale. Wheeless and Grotz's (1977) ITS is a unidimensional scale measuring one's trust for another specific individual. The scale consists of 15 semantic differential-type, 7-interval items with the positive-negative

polarities randomly ordered to avoid response bias. A split-half reliability coefficient of 0.92 was reported when the scale was administered to 100 teacher and their spouses or eldest children ($n = 261$). The authors report that “the instrument had predictive validity ... for general use as an alternate means of measuring trust” (p. 256).

Our Class and Its Work. Eash and Waxman’s (1983) OCIW instrument was used for gathering student perceptions of teacher classroom behavior. The instrument consists of 40 items describing teaching behaviors that form eight Likert-type subscales, namely, didactic instruction, enthusiasm, feedback, instructional time, opportunity to learn, pacing, structuring comments, and task orientation. High reliabilities for the eight scales are reported; after administering the scale to 762 pupils in 36 grade 4 and 6 classes in a large American city school district Cronbach alphas range from 0.84 to 0.92. To assess construct validity, an examination of the relationships among OCIW scores, student achievement, and principal ratings of the teachers obtained from the sample described showed that teachers who obtained higher scores on the OCIW scale also “were rated higher by the principal in his yearly evaluations” (p. 4).

Pupil achievement grades. This study relied entirely on the teachers’ anecdotal and letter graded report cards for establishing levels of pupil achievement. Achievement data from teachers’ anecdotal notes on report cards of pupil progress were used to check the veracity of the teachers’ assigned behavior grades. Two report cards of pupil progress were used in this study: (a) November 1991—premeasure, and (b) June 1992—postmeasure. Pupil report card data (letter grades) regarding achievement (e.g., language arts, math, science, and social studies were equally weighted to determine an overall achievement score) were coded using a four-point grade point average (GPA) type scale (A+, A, A– were assigned a 3, B+, B, B– were assigned a 2, C+, C, C– were assigned a 1, D+, D, D–, F were assigned a 0).

Pupil behavior grades. Pupil behavior data were collected from the November 1991—premeasure—and the June 1992—postmeasure report cards. Behavior data from teachers’ anecdotal notes on report cards of pupil progress were used to check the veracity of the teachers’ assigned behavior grades. Categorical data were converted to numerical data using the following schema: (a) good or excellent—assigned a 2, (b) satisfactory—assigned a 1, and (c) poor or needs improvement—assigned a 0. These behavior grades were then averaged to obtain an overall class behavior grade.

School Attitudes Scale for Children. The SASC consists of a series of “semantic differential scales that tap the following school-related dimensions: School, Teachers, Arithmetic, Science, Social Studies, Language Arts, Music, Drama, French, Art, Dance, Religion, Health, and Physical Education” (Randhawa & Van Hesteren, 1982, p. 6). The authors reported reliability coefficients ranging from 0.89 to 0.97 for the scales when administered to 99 pupils in grades 3-6. The authors also report that the scales exhibit a high degree of construct validity when analyzed using a “multitrait-multimethod convergent and discriminant validation methodology” (p. 11). For the present study four of the original SASC scales were used: school, social studies, language arts, and arithmetic. In addition to these, four other SASC scales—yourself, your

classmates, your regular teacher, and learning in general—were created to obtain measures of what the literature suggests are important attitudinal objects.

Supervisory Beliefs Inventory. The SBI (Glickman, 1990) is used to estimate the proportion of preferred modes of interaction—directive, collaborative, and nondirective—between a teacher and a teaching partner. The SBI consists of 15 items that when scored indicate the approximate proportion of preference for each of the interaction approaches. The literature reveals no reliability measures for the SBI. With respect to validity, the author states that “the instrument has been field-tested six times with ninety supervisors and supervisor trainees. Response between the options indicated ‘good’” item discrimination” (p. 91).

Teacher Efficacy Scale. Gibson and Dembo’s (1984) Likert-type teacher efficacy scale contains two factors: general teaching efficacy, the belief that teachers as a collective can influence pupil learning; and personal teaching efficacy, the belief that the individual teacher can influence pupil learning. For this study the differentiation between general teaching efficacy and personal teaching efficacy was maintained. The authors report the following Cronbach’s alpha coefficients for the 16-item teacher efficacy scale: 0.78 for the personal teaching dimension, 0.75 for the general teaching dimension. These findings are supported by Anderson, Greene, and Loewen (1987), who reported similar findings.

Judith Golec

John Gartrell

University of Alberta

and

William J. Sveinson

University of Victoria

University Performance of Nonmatriculated Admissions

Nonmatriculated admission (NMA) programs in universities have sought to promote access and equity while maintaining academic standards. These programs often assume that maturity is associated with improved motivation, interest, and decision making abilities and that life experience will provide the kind of contextual information that improves learning. These assumptions appear to be overly optimistic in their expectations. The academic performance of mature NMA admissions to the Faculty of Arts at the University of Alberta was in fact substantially poorer than the performance of matriculant admissions. The problem with this admission program may be attributable to abandonment of individual screening in favor of minimalist, formal admission criteria.

Les programmes d'admission aux universités pour les élèves non diplômés du cours secondaire 2^e cycle ont cherché à promouvoir l'accès et l'équité tout en maintenant certains standards académiques. Ces programmes présument souvent qu'à la maturité s'associe un accroissement de motivation, d'intérêt, et d'habiletés à prendre des décisions et que les expériences de vie fourniront l'information contextuelle nécessaire à l'amélioration de l'apprentissage. Ces suppositions sembleraient être trop optimistes dans leurs attentes. La performance académiques des élèves adultes non diplômés du cours secondaire 2^e cycle à la faculté des arts de l'Université de l'Alberta était en effet considérablement inférieure à celle des admissions des élèves diplômés du cours secondaires 2^e cycle. La difficulté que présente ce programme d'admission peut être attribuée à l'abandon des interviews individuelles des candidat(e)s en faveur des critères d'admission formels minimalisés.

Nonmatriculated admission (NMA) programs have been a part of Canadian universities' efforts to promote access and increase enrollments. More relaxed admission policies represent a response to equity issues, because mature admissions give people a "second chance." To the extent that universities become centers of adult retraining and learning over the whole life course, it may be

Judith Golec is an assistant professor in the Department of Sociology. She is currently studying the process of mutual aid in health-related social support groups.

John Gartrell is a professor in the Department of Sociology. He is reanalyzing the data used for *The Bell Curve* and beginning a study of development and the environment in Canada's Rocky Mountain communities.

William Sveinson is a research project coordinator at the Centre on Aging. He is presently providing management and technical support for a major study of independent community living among seniors.

seen as unfair to exclude a large number of people (Alberta Advanced Education, 1990; Anisef, Okihiro, & James, 1982; Vaselenak, 1970). The problem with flexible admission requirements is the maintenance of academic standards. Of course, the admission of nonmatriculated mature applicants assumes that previous failure can be attributed in large part to immaturity rather than inability or lack of academic preparation. The gaining of life experience that comes with greater maturity is expected to provide prospective students with the ability to overcome previous difficulties. For this reason nonmatriculated admission is sometimes called mature student admission.

Our focus is on standards of student performance. More specifically, do mature admissions to university perform as well as students who are admitted with full high school matriculation? The assumption of mature admission programs is that experience can compensate for prior poor performance. Experience is thought to provide increased commitment, motivation, interest, and decision making abilities, which make up for a lack of formal academic qualifications. Knowledge gained through life experience is thought to provide an enriched context for the interpretation of new knowledge (Eaton & West, 1980; Jackson, Small, & Zelmer, 1985; Perkins, 1971). Unfortunately, the efficacy of nonmatriculation admission policies is often just taken as a matter of faith. Professors may assume that mature students do better because they have had some mature students who were outstanding (overgeneralization), but there is remarkably little systematic, published evidence to support this article of faith.

In Canada, Darling (1985) documented the extent and diversity of mature admission programs. All 44 postsecondary institutions surveyed had special provisions to admit nonmatriculated applicants. Perkins (1971) concluded that mature admissions in Arts and Humanities at the University of Lethbridge were superior in their overall academic performance. "Maturity and motivation compensated for lack of past formal educational experience" (p. 111). Jackson et al. (1985) assert, on the basis of nonsystematic observation, that NMA performance at the University of Alberta was satisfactory. Wilson and Lapinski (1978) found similar results in a more systematic study conducted at Simon Fraser University, although a later unpublished report (Mature Student Entry Advisory Committee, 1983) comes to the opposite conclusion. There is some corroboration of the success of NMA programs from studies conducted in Australia (Barrett & Powell, 1980; Eaton & West, 1980) and Great Britain (Smithers & Griffin, 1986; Walker, 1975). More generally, a number of authors have documented a positive correlation between age and learning, even among adults (Johnson & Walberg, 1989; McLeish, 1962). As sparse as the empirical evidence appears to be, it seems to offer support for an organizational culture in universities that believes that mature students perform better than, or at least as well as, students with high school matriculation.

Admission policies of the Faculty of Arts at the University of Alberta provide an opportunity to empirically examine the academic performance of NMAs. In many respects the U of A's experience of relaxing normal admission requirements is not unusual. The earliest examples occurred during the 1940s. In the early 1940s the high school curriculum for the province as a whole was revised. However, not all secondary schools, especially rural schools, were able to offer the full high school curriculum. At the discretion of the dean of the

admitting faculty, some applicants with one or more matriculation deficiencies were admitted in the belief that returning to high school for upgrading was impractical. This is the same logic that was used a few years later to admit ex-servicemen whose education had been interrupted by World War II. It was also at this time that the University of Alberta began to offer matriculation (100 level) subjects as part of its own curriculum, thereby ensuring that deficiencies would be remedied during the first year of university.

Like most universities in Canada (Darling, 1985) it was not until the 1960s that the University of Alberta created a formal policy for admitting unqualified "adults" as part of a broader movement to increase access for students who had dropped out of formal education, the late bloomers, and the economically disadvantaged. The small numbers of applicants admitted to the Faculty of Arts (7 in 1965-1966; 31 in 1966-1967; and fewer than 65 in any year until 1973-1974) suggests a cautious application of the policy. In fact, unqualified applicants were required to undergo a rigorous screening process that included an initial interview with the Assistant Dean, a battery of tests administered by Student Counselling Services, an essay explaining the applicant's work experience and reason for desiring to attend university, followed by another interview with the Assistant Dean. Because decisions were based on the merits of the case, many students were advised to take remedial studies before reapplying. Interestingly, applicants lacking only one matriculation subject were regularly required to complete matriculation. Apparently, it was felt that taking one subject did not entail much hardship.

Between 1972-1973 and 1973-1974, enrollment of nonmatriculants jumped markedly (203) and in-house faculty reports document a concern that academic standards were being diluted. During the 1970s and early 1980s the non-matriculated admission policy and procedures underwent many changes. Regrettably, because the experiments were not formally evaluated, it is not possible to state what impact the various changes may have produced. The minimum age was changed from 21 years to 24 and back to 21. For a short period NMAs were admitted "on probation." A minimum level of academic preparation was prescribed (grade 10) and later removed. Citizenship (Canadian or landed immigrant) and residency (Province of Alberta, Northwest Territories, or the Yukon) requirements were specified. Applicants lacking a matriculation level language other than English were admitted but on the condition that they clear the deficiency during the first year of the program; through practice this condition was extended with completion required sometime before graduation. Administering the battery of aptitude tests exacted a relatively high cost in organizational resources, and besides, very few applicants were rejected on these grounds alone. Screening applicants with a battery of psychological tests was reinterpreted as a barrier to access and the practice ceased. Similarly, the practice of requiring essays or letters of intent was also abandoned.

By the mid-1980s, the only admission criterion for NMAs besides age (21 years), citizenship, and residency was the successful completion of grade 12 English (English 30) or its equivalent (writing competence test or a university level English). If other academic work had been attempted in the three years preceding the application for mature admission, there could be no record of

failure. Apparently there is nothing unique about all this. According to the Darling (1985) survey of nonmatriculation admission policies, of the postsecondary institutions prescribing a minimum age (only 10 do not) the majority (23 out of 34) specify 21 years. Like the University of Alberta, roughly half (21 of 44) grant NMAs clear rather than probationary admission and most (33 of 44) use the same procedures to admit to full- or part-time status. Most (all but 6) have either ceased or never introduced the practice of screening NMAs with tests. Unlike the University of Alberta, about half still require a letter or résumé to evaluate the applicant's life experience and reason for entering a degree program.

Enrollment of NMAs increased. By 1985-1986 they comprised roughly 40% of freshman undergraduate admissions to the Faculty of Arts. In the absence of reliable data that describe enrollment trends for the NMA category, it is not possible to judge with certainty whether the Faculty of Arts at the University of Alberta is unique in this regard. Relying on impressions of the registrars surveyed, Darling (1985) estimates that NMA admissions make up fewer than 5% of all full-time undergraduate admissions and more for part-time admissions (some faculties in excess of 10%).

In the face of declining budgets, increasing demand for undergraduate admission, and a stable organizational resource base, the University of Alberta (like other universities) was forced to consider strategies for controlling undergraduate enrollments. In fact during the mid-1980s, the minimum admission average for matriculated applicants was raised from 65% to 70%, and discussion on the merits of establishing enrollment quotas in the undergraduate faculties was initiated. There is an obvious contradiction contained in these admission policies: increasing academic standards to reduce the numbers of fully qualified applicants while relaxing standards to increase the numbers of unqualified mature students. This puts a new turn on the issue of access and provides the social context for examining the academic performance of NMAs.

Our investigation was guided by a number of general propositions. First, based on previous findings and the admission policies for the Faculty of Arts and the University of Alberta, we would expect NMAs to perform at least as well as matriculation students. Second, to the extent that age reflects maturity, which in turn enhances performance (perhaps via experience and motivation), we expect age to have a positive effect on performance for both regular matriculation and NMA admissions. Third, we wish to examine the accuracy with which other university admission criteria (English 30 grade) predict academic performance. Finally, we extend our analysis to investigate how early performance in university predicts later performance, particularly after adjusting for the effects of age and admission criteria.

Research Design

Sample

All undergraduate students admitted (for the first time) to year 1 of a three- or four-year degree program in the Faculty of Arts between September 1985 and January 1986 were selected for study. This includes students applying under matriculation and nonmatriculation status.¹ The academic performance of the 1985-1986 study cohort was observed over a four-year period ending in August 1989.

The data for this analysis were collected entirely from student records (application form, academic transcripts) and therefore are limited to information routinely collected and held by the Registrar in electronic form. One advantage of using student records is the cost efficiency of available information. The use of records data also makes it likely that our analysis could be easily replicated at other universities or colleges. However, the University of Alberta student records were not designed for research use, and it was therefore necessary to make data transformations (described below under "measurement") in order to carry out the analysis. A more serious drawback to archival data is the limitation imposed by the paucity of information contained in such records. The records do not contain information on employment history, family commitments, or other variables that might help to shed light on questions of life experience, motivation, and maturity.²

Measurement

We selected grade point average (GPA) as our primary indicator of academic performance. GPA is constructed by dividing each student's total grade points by his or her total course weights taken in a specified time period. GPA has a number of advantages over other performance criteria, because it is commonly used by universities and the students themselves to evaluate student performance. As a composite of many professors' assessments of the students' performance, random errors in assessment are more likely to average out. At the University of Alberta course grades (and grade point averages) are measured on a 9-point scale (1-9). A score of 3 represents a failure, 4 is a marginal pass, 5 and 6 are satisfactory (roughly high and low Cs), 7 represents good performance (a B), 8 is very good or excellent (an A), and 9 is outstanding (an A+).

Sessional GPAs are calculated and printed on official transcripts at the University of Alberta. Traditionally, most students attended only winter session (September through April), and academic decisions for promotion and distinction continue to be based on the winter session GPA. In the case of matriculants, winter session GPA still tends to capture the yearly course load for most students. In the case of nonmatriculants, however, yearly course work often stretches across three sessions: winter, spring, and summer. To avoid distortion, we measured academic performance by computing an overall (four-year) GPA, that is, by summing grade points and course weights across the three sessions and four years under study. We were also interested in examining the effect of early (first-year) performance on later performance (years 2 through 4). Therefore, we computed a first-year GPA by summing grade points and course weights across the three sessions of 1985-1986 and a measure of later performance by computing a GPA for years 2 through 4 (summing across all sessions for 1986-1987, 1987-1988, and 1988-1989).

Other indicators of academic performance are also reported. These include graduation, first-class standing, receiving a scholarly award, and the requirement to withdraw. We compare the percentage of matriculants and nonmatriculants who graduated on or before spring convocation 1989, earned first-class standing (sessional GPA of 7.5 or higher on a full course load) in any winter session, and received at least one scholarship or prize in four years. Because nonmatriculants are more likely than matriculants to attend university as part-time students, their part-time status (if not their GPA) more often

makes them ineligible for first-class standing and scholarly awards. Understandably, part-time attendance also delays graduation. The requirement to withdraw for academic reasons is our final indicator of academic performance.

Annual decisions concerning withdrawal (and promotion) are made every spring at the completion of winter session. These decisions are based on a minimum of nine course weights (three term courses) taken during winter session. They require the calculation of a quality index (QI). The QI is calculated by dividing the sum of course weights passed by the sum of course weights attempted and multiplying the result by the sessional GPA. The QI, like the GPA, ranges from 1 to 9. A QI of 2.0 or less in a single winter session or a QI between 2 and 4 in any two winter sessions results in a requirement to withdraw. In spite of poor performance, part-time students completing only nine course weights in winter session escape compulsory withdrawal and avoid having the notation "required to withdraw" permanently imprinted on their academic transcript.

The primary independent variable is admission status: admission on the basis of high school matriculation (coded as 0) or NMA standing (coded as 1). Our primary objective is to investigate differences between matriculants and NMAs in their performance at university. In a sense we want to investigate whether the maturity (older age) of NMAs is in some sense a viable substitute for matriculation.

Age is also an important variable. It is measured as age in years on admission. For September admission, students must be 21 years of age before September 30, and for January admission (of which there were few), the 21st birthday must occur before January 30. However, age is confounded by admission status, because age is one of the principle criteria for admission as an NMA. All NMAs were 21 years of age or older, and all but 25 of the 579 matriculant admissions were under 21. Still, if maturity has a pervasive effect on performance, age differences may have a positive effect on GPA in admission cohorts. A year or two can mean a great deal to the matriculant admissions, and NMA admissions in their late 20s and beyond may be considerably more mature than those in their early 20s. We have to be aware of the close relation between age and admission status when interpreting our results.

Past academic performance is also obviously relevant to performance at university. However, we were limited in our selection of indicators, because NMAs by definition lacked relevant high school records that were comparable to matriculation admissions. Besides age (and a lack of recent academic failure), the only other criterion for admission of NMAs in 1985-1986 was grade 12 English (English 30). It was therefore used to examine the effects of the only indicator of prior performance used in the admission decisions. Matriculation admissions were scored on the basis of the percentage grade they received in high school English as reported on secondary school transcripts. Non-matriculant admissions were also given their percentage grade in high school English if they had taken the course. The vast majority of NMAs (approximately 90%) satisfied the English criterion with a high school English. About 9% of the NMAs satisfied the English admission criterion by completing the Writing Competence Test. Grades on this test were converted to percentages using the formula devised by the Registrar's Office.³ Only four students com-

pleted a university level English course. If NMAs satisfied the admission English criterion by completing a first-year university level English course, their grades on the 9-point scale were reconverted to percentages using the University's conversion formula.⁴

We also were able to add a variable representing whether or not the student had a deficiency in the requirement for a language other than English (LOE) on admission (a deficiency is coded as 1; no deficiency is coded as 0). This information was recorded as part of the admission ruling for this and earlier cohorts because all Arts students, as a condition of graduation, were required to present credit in a language other than English either at the senior matriculation or university level. Since 1990, the LOE has been introduced as an admission requirement at the senior matriculation level and as a degree program requirement at the university level. LOE deficiency, therefore, gives us a second crude indicator of prior performance that has additional policy relevance. Presumably, the assumption is that if a mature student has an LOE credit, they are more likely to perform better in the Faculty of Arts.

Results

Admission Status and Overall GPA

Our analysis focuses on the 579 matriculants and 332 NMAs admitted to the Faculty of Arts during the 1985-1986 academic year.⁵ Contrary to our predictions, the 332 NMAs did not perform better at university than the matriculants (Table 1). The NMAs' overall GPA was 1.06 lower than that of the matriculants. The simple regression of GPA at the end of four years on type of admission (with admission status coded NMA=1, Matric=0) reconstructs these same results a little differently. The intercept was 5.97 (the mean for the Matrics) and the slope was -1.06 (the mean for the NMAs was 1.06 lower). This simple admission status difference accounted for 7.6% of the variance in GPA.

Table 1
Student Characteristics by Type of Admission

Student Characteristics	Total Sample	High School Sample	Mature Admission (NMA)
Average Age (Yrs.)	22.2	18.3	27.4
		Percent	
Female	60	61	57
English 30 grade	71	74	65
LOE deficiency	49	28	84
Won Award	5.4	7.1	2.4
Graduated	33	40	19
First-class Standing	16	21	8
Ever forced to withdraw	12	12	13
Overall GPA	5.59	5.97	4.91
Std. Dev.	(1.87)	(1.52)	(2.21)
Sample Size	911	579	332

Other indicators of overall performance are consistent with these results. This is expected because the requirement to withdraw as well as the achievement of awards, first-class standing, and graduation are partly dependent on GPA. By the end of four years, matriculants were more likely to have won awards (7.1% vs. 2.4%), to have achieved first-class standing in any year (21% vs. 8%), and to have graduated (40% vs. 19%). High school matriculants were only slightly less likely to have ever been forced to withdraw (12% vs. 13%).⁶ On the average, matriculants had only slightly better English 30 marks than NMAs (74% vs. 65%), although as we would expect, they were much less likely to have been admitted with an LOE deficiency (28% vs. 84%).

The Effects of Age and Sex

When we add the demographic factors of age and sex to the regression (Table 2) we increase the R-square to .110. For the whole entering cohort, age was virtually uncorrelated with GPA (Table 2, $r = -.065$). Older students did not have higher GPAs. However, when we controlled for admission status we did observe the predicted positive effects of age on GPA (St. Beta=.217). Admission status suppressed the positive effects of age on GPA. Controlling for gender and admission status, for each year increase in age, predicted GPA increased by $b = .063$. For a mature student entering in his or her early 30s instead of early 20s (10 years difference in “maturity”), we would predict a GPA about .63 higher.

The observed zero-order effects of gender on GPA were small and negative ($r = -.109$). This weak tendency for males to have somewhat lower GPAs than females was not greatly altered with the addition of controls for age and admission status (St. Beta=-.087). GPA for females averaged only slightly higher than males (about .331 grade points), controlling for admission status and age. This confirmed our expectation that there would be only small gender differences in GPA. However, when age and gender were controlled, the differences between matriculant and NMA admissions increased to 1.626 from the observed difference of 1.060 and the zero-order correlation ($r = -.275$) increased to a standardized beta of $-.419$. This suppressor effect was produced largely by the strong positive correlation ($r = .679$) between age and admission status. As we expected, age and admission status were partially confounded,

Table 2
Overall Grade Point Average: Admission Status, Age, and Sex

Variable	b	St. Error	St. Beta	Tolerance	r
<i>Regression Results</i>					
Admission Status (NMA=1, Matric=0)	-1.626	.166	-.419	.536	-.275
Age (years)	.063	.012	.217	.536	-.065
Sex (Male=1, Female=0)	-.331	.120	-.087	.994	-.109
Constant	4.939	.248			

$R^2 = .110$; $df = 3, 906$.

because age was used as an admission criterion for NMAs. However, despite the strong correlation between age and admission status, there was no evidence of substantial collinearity. Tolerance remained above .5 for both variables (Table 2), and standard errors for the partial correlation coefficients were not inflated. Further tests revealed no significant interaction effects among age, gender, and admission status as they affected the overall GPA.⁷

Preuniversity Performance and GPA

To see if preuniversity academic performance accounted for the performance difference of Matrics and NMAs we added English 30 grades and LOE deficiency to the regression equation (Table 3). We expected both to have significant effects on GPA, and we also expected them to account for part of the effects of admission status.

Adding indicators of prior performance increased the R-square to .184. When we controlled for the two indicators of previous academic performance, the effect of type of admission was reduced to .892 points. Prior academic performance explained a little less than one half of the admission status differences observed in Table 2. Still, admission status had substantial effects on GPA (St. Beta=−.230). Gender effects now had virtually no effect on GPA (St. Beta=−.032). The effect of age remained relatively unchanged (b=.060; St. Beta=.205). LOE deficiency on admission had little effect on GPA (St. Beta=−.090). Those without a language other than English had average GPAs that were .337 lower. English 30 grades had a larger partial effect (b=.056; St. Beta=.288). A difference of 10 points on a student’s English 30 grade would predict a GPA that was .56 points higher.

First-year Performance and GPA

Finally, in order to examine the effects of early performance at university on later GPA, we recomputed our dependent variable to represent average GPA for courses taken during years two through four. We then included the average of grades obtained in the students’ first year (1985-1986) as a predictor. As we would expect, the R-square for the equation increased considerably (to .400).

Table 3
Admission Status and Overall GPA

Variable	b	St. Error	St. Beta	r
<i>Regression Results</i>				
Admission Status (NMA=1, Matric=0)	−.892	.185	−.230	−.275
Age (years)	.060	.012	.205	−.065
Sex (Male=1, Female=0)	−.122	.119	−.032	−.109
English 30 Grade	.056	.007	.288	.374
LOE Deficiency	−.337	.138	−.090	−.252
Constant	.874	.554		

R² = .184 df = 5, 904.

First-year grades were strongly correlated with GPA for years two through four ($r=.602$), and first-year GPA was an excellent predictor of later GPA (St. Beta=.576; $b=.938$).⁸ When first-year grades were controlled, the effect of admission status was reduced somewhat ($b=-.711$). Still, the effect of admission status persisted throughout years two through four even when we controlled for initial performance.

When first-year performance was controlled, English 30 grades had little effect on subsequent university grades (St. Beta=-.032). LOE deficiency had a somewhat larger effect (St. Beta=-.098), but this effect was still small. These results suggest that any effects of these two factors was largely captured by first-year GPA. The effect of first-year grades also appeared to account for most of the observed effect of age on later GPA (St. Beta=.032). Controlling for first-year GPA (and the other variables in Table 4), the partial regression coefficient for age was reduced to almost zero ($b=.008$, compared with $b=.060$ in Table 3). The greater maturity of older students helps them initially, but it does not appear to help them more as they progress through their university studies.⁹

Discussion

Contrary to our expectations and those of the University of Alberta, the NMAs from this cohort performed substantially more poorly than Matriculation admissions. At the end of four years, the less qualified mature students averaged a full grade point less than Matric admissions. They still scored an average of .89 grade points less when overall GPA was adjusted for demographic and prior performance factors.

Gender differences were negligible. By contrast, age differences in performance at university were substantial. Yet age had little effect on later performance in years 2 through 4. It therefore appeared that the principal effect of age differences occurred early in a student's university career. English 30 grades had similar effects. They contributed to the prediction of overall GPA, but when GPA for years two through four was regressed on first-year GPA,

Table 4
Determinants of GPA for Years 2-4

Variable	<i>b</i>	St. Error	St. Beta	<i>r</i>
<i>Regression Results</i>				
First Year GPA	.938	.04	.576	.602
Admission Status (NMA=1, Matric=0)	-.711	.237	-.125	-.278
Age (years)	.008	.015	.032	-.066
Sex (Male=1, Female=0)	.117	.009	-.032	.267
English 30 Grade	-.009	.009	-.032	.267
LOE Deficiency	-.533	.177	-.098	-.267
Constant	.929	.708		

$R^2 = .400$ $df = 6, 868$.

English 30 grades had little impact. The effect of English 30 appeared to be “captured” by first-year performance. Whether or not the student had a deficiency in a language other than English (LOE) had little effect on GPA.

What implications do these results have for formulating policy with respect to the admission of NMAs? The consistently lower averages of NMA students compared with regular Matric admissions are persuasive. Whatever maturity is associated with NMA status apparently does not make up for lack of qualifications. However, age did have positive effects on initial performance (first-year GPA) and NMA students are generally older than Matric admissions. In order to make up for this observed GPA deficit, NMAs would have to be about thirty years old when admitted. From this perspective, the age criterion for NMA admission would have to be increased substantially.

In 1990 the Faculty of Arts did raise the minimum age for NMA admission from 21 to 24 years. This was part of a more restrictive move to manage enrollments while maintaining academic standards (Committee on Admissions and Transfer, 1988). A closer examination of the 1985-1986 study cohort, by age groups and overall GPA, is instructive (Table 5). For the Matric students, the youngest admissions (16-17 years) have a higher overall GPA than any of the age groups except for those few students who were 24 years or over on admission. This apparently curvilinear effect of age on GPA for Matric admissions may have been something of an historical anomaly. Cohorts entering at age 18 through 23 contain a substantial proportion of the increasing numbers of students who during the 1980s took more than one year to complete grade 12 (Alberta Advanced Education, 1987). This includes students who went back to get particular subjects, but it also includes a substantial proportion of students who retook courses in order to raise their average. The difficulties that these students experienced in high school may simply continue with poorer performance in university. It is not until matric students reach the age of 24 years that their performance exceeds that of students who completed high school in the standard minimum time (16- to 17-year-olds).

Table 5
Age and Overall GPA by Admission Status

Age (years)	Matric	N	Nonmatric	N
Overall GPA				
16-17	6.22	130	*	
18	5.94	344	*	
19	5.67	61	*	
20-23	5.45	28	4.33	127
24-29	6.59	9	4.99	111
30-39	7.03	6	5.37	68
40 and over	6.00	1	6.22	26
Total	5.97	579	4.91	332

*All NMAs were 21 years of age or older.

Table 6
The Effects of Age on Overall GPA: One-way ANOVA for Matric and Nonmatric Admissions

	Admission Status	
	Matric	Nonmatric
F	2.32	7.51
Sig. F	.032	.000
Eta ²	.024	.064

For the nonmatric admissions, only those who are 40 years of age or older had higher overall GPAs than any of the matric admission age groups. The youngest nonmatric admissions, who constituted over one third of the nonmatric cohort, had by far the poorest performance. Based on our results, we would expect the change in prescribed age for NMA admission to improve GPA although somewhat less than was hoped for. It appears to take a good deal of maturity and experience, at least as reflected in chronological age, to compensate for lack of academic preparation. Still, age did improve performance for both matric and particularly nonmatric admissions (Table 6).

The Faculty of Arts introduced another change in 1990. The language other than English was prescribed as an entrance requirement for NMAs. Fulfillment of this requirement preadmission would obviously facilitate the completion of the Arts degree language requirement. However, this additional entrance requirement (in combination with a higher minimum age) had the effect of reducing NMA admissions to a comparative trickle (51 in 1990-1991, 55 in 1991-1992, 49 in 1992-1993). Based on our results, we would not expect this additional entrance requirement to improve subsequent university performance in other subjects. LOE deficiency had little effect on GPA.

Why did the Faculty of Arts at the University of Alberta not get the academic performance of nonmatriculated admissions that previous studies would have predicted? We can only speculate, but the answer seems to lie in the reflexive experience that screening procedures make available. The successful NMA programs in Canada (Perkins, 1971; Wilson & Lapinski, 1978), Australia (Barrett & Powell, 1980; Eaton & West, 1980) and Britain (Smithers & Griffin, 1986; Walker, 1975) all employed rigorous methods of personalized screening as part of the admission process. It may not matter if the entrance test (whether this is a critical essay or an exam taken after a compulsory lecture on an unknown topic) is even scored (Barrett & Powell, 1980). What matters more is that the method used should acquaint "mature" applicants with the nature and demands of university study and supply them with the chance to assess their own abilities, interests, and motivation.

A higher participation rate in university education is one of the most important achievements in the past three decades. Bureaucratization and standardized decision making accompanied the shift to mass education in postsecondary institutions. The University of Alberta was not alone in abandoning labor-intensive individualized screening of NMA admissions in favor of minimalist formal criteria. The lesson may be that access and academic

performance have to be carefully balanced. Increased access for mature students cannot be successfully implemented without the application of significant organizational resources.

Notes

1. Besides "regular" NMAs, the Faculty of Arts also admits applicants without matriculation who (a) were previously admitted to another university on the basis of mature standing or (b) had successfully completed the first year of a technical or vocational program in Alberta. These two admission categories include only a relatively small number of freshman students (51 and 67 respectively in 1985-1986) and they are not necessarily "mature" (21 years or older). Both categories are omitted from the study.
2. Cost considerations precluded collecting supplementary information directly from students.
3. There were only two passing grades: S (satisfactory) was scored as 85%; MS (mostly satisfactory) was scored as 65%. If students received an MU (marginally unsatisfactory) or U (unsatisfactory), they would not satisfy the admission English criterion.
4. The conversion formula assigns each stanine score the midpoint of the percentage range to which it refers: 9=95, 8=85, 7=76, 6=69, 5=62, 4=54. Grades below 4 would not satisfy the admission criterion.
5. Interpretation of results emphasizes the magnitude of effects, because our sample was not designed to be directly representative of a larger population. Tests of significance are reported, but they are less important than the magnitude of the effects observed (Cohen, 1991).
6. Closer examination of the record reveals that a large part of the reason for this particular finding is that nonmatriculants are more likely to withdraw voluntarily. An indication of this is that more NMAs (17% vs. .04%) actually completed fewer than three courses in four years. All but five of the students falling in this category failed to accumulate any grade points after the first year.
7. Given the policy relevance of age 21 (admission criterion for NMA in this cohort) and 24 (the new 1990 criterion), we examined separate regressions of sex and admission status on GPA for both of these age splits. No significant differences in effects were observed.
8. We might note that there is no evidence of grade inflation in these results. Had there been substantial grade inflation from the first year to subsequent years, we would have expected to observe a partial slope for first-year GPA that was greater than 1.0.
9. There was no significant interaction between admission status and the effects of initial performance on later performance.

Acknowledgments

This research was supported in part by Student Temporary Employment Program Grant LPT 216576. The authors gratefully acknowledge the assistance of Vic Stobee in the Faculty of Arts and Marj McMullen in the Office of the Registrar. We also thank an anonymous reviewer for helpful comments.

References

- Alberta Advanced Education. (1987). *Enrolment and population trends: A background paper*. Edmonton, AB: Research and Information Support.
- Alberta Advanced Education. (1990). *Responding to existing and emerging demands for university education: A policy framework*. Edmonton, AB: Author.
- Anisef, P., Okihiro, N., & James, C. (1982). *Winners and losers*. Toronto, ON: Butterworth.
- Barrett, E., & Powell, J.P. (1980). Mature age unmatriculated students and the justification of a more liberal admissions policy. *Higher Education*, 9, 365-383.
- Cohen, J. (1991). Things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- Committee on Admissions and Transfer. (1988). *Report of the subcommittee on non-matriculated admissions*. Edmonton, AB: University of Alberta.
- Darling, A.L. (1985). The admission of non-matriculated adults to university credit courses. *Canadian Journal of Higher Education*, 15(2), 91-97.
- Eaton, E., & West, L. (1980). Academic performance of mature age students: Recent research in Australia. In T. Hore & L. West (Eds.), *Mature age students in Australian higher education* (pp. 48-68). Victoria, Australia: Monash University.
- Jackson, L.S., Small, J.M., & Zelmer, A.E. (1985). Mature student provisions at the University of Alberta. *Canadian Journal of University Continuing Education*, 11(2), 64-74.

- Johnson, M.L., & Walberg, H.J. (1989) Factors influencing grade point averages at a community college. *Community College Review*, 16(4), 51-60.
- The Mature Student Entry Advisory Committee. (1983). *Final report*. Burnaby, BC: Office of the Registrar, Simon Fraser University.
- McLeish, J. (1962). The age factor in adult learning. *International Review of Adult and Youth Education*, 14, 26-45.
- Perkins, S. (1971). Success of mature age students in university studies. *Alberta Journal of Educational Research*, 17, 105-112.
- Smithers, A., & Griffin, A. (1986). Mature students at university: Entry, experience and outcomes. *Studies in Higher Education*, 11(3), 257-268.
- Vaselenak, M. (1970). Admission of mature non-matriculated students into a degree program. *McGill Journal of Education*, 5(1), 85-96.
- Walker, P. (1975). The university performance of mature students. *Research in Education*, 14, 1-13.
- Wilson, L., & Lapinski, B. (1978). A longitudinal study of mature students in credit programs. *Canadian Journal of Continuing Education*, 4(2), 22-27.

June Beynon

and

Kelleen Toohey

Simon Fraser University

Access and Aspirations: Careers in Teaching as seen by Canadian University Students of Chinese and Punjabi-Sikh Ancestry

This study, based on interviews with 34 Canadian university students, 22 of Chinese ancestry and 12 of Punjabi-Sikh ancestry, revealed that cultural influences (i.e., parents' views about the desirability of teaching) and structural barriers (i.e., discrimination in schools and insufficient instructional support in academic English) are implicated in individuals' perceptions of teaching careers. Increasing participation in teaching requires a theoretical perspective that includes these factors as well as the views of individuals who interpret and react to them. This research shows there is diversity in the way these factors operate in the two ethnic groups, as well as diversity based on gender and individual perspectives and experiences. Recommendations made for increasing participation in teaching attend to this variability.

Cette étude basée sur des interviews accordées à 34 étudiant(e)s universitaires canadien(ne)s dont 22 d'ascendance chinoise et 12 d'ascendance punjabi-sikh, révèle que les influences culturelles (ex., l'opinion des parents concernant l'attrait à l'enseignement et les barrières structurelles (ex., la discrimination dans les écoles et l'insuffisance du support dans l'enseignement de la langue anglaise) sont tenues en ligne de compte dans leur perception de l'enseignement comme choix de carrière personnel. La participation croissante dans l'enseignement requiert une perspective théorique qui doit inclure ces facteurs ainsi que les opinions des individus qui les interprètent et qui y réagissent. Cette recherche démontre qu'il y a une diversité dans la façon qu'opèrent ces facteurs dans ces deux différents groupes ethniques ainsi qu'une diversité fondée sur le sexe, et les points de vue et les perspectives et les expériences individuels. On recommande d'accroître la participation dans l'enseignement pour répondre à cette variabilité.

Introduction

Having minority teachers would definitely help minority students in adapting to Canadian culture. Minority teachers would be more aware of the problems students might be having and how to help them. (Peter, Canadian university student of Chinese ancestry)

June Beynon is an associate professor in the Faculty of Education where she teaches courses in multicultural and antiracist education. She has also worked as the multicultural education officer for the Vancouver School Board. She is currently researching First Nations teachers' perceptions of their roles as social change agents.

Kelleen Toohey is an associate professor in the Faculty of Education where she teaches ESL methodology. Currently she is conducting research in the integration of second language learners into Anglophone classrooms.

To respond to the needs of minority students, as Peter has outlined, and to address concerns about equity of access and equity of representation in employment are the prime reasons that educational practitioners and policy makers are currently advocating recruitment and training of minority teachers.¹ In addition, having teachers who are representative of the social, cultural, and linguistic diversity of Canadian society provides the possibility for all students to learn from individuals who represent culturally diverse perspectives. In our role as teacher educators, we feel obliged to discover the character of barriers at the university that could prevent minority group individuals from entering teacher education programs. Students' perceptions of the desirability of teaching as a career, as well as their views of the particular factors that enhance or inhibit access to this career, may be useful to us and potentially other institutions in understanding the reasons for current underrepresentation and what we might do to address the problems.

The purpose of this study was to explore the perceptions and articulations of experiences affecting career choices made by Canadian university students of Chinese and Punjabi ancestry (the two most populous visible minority communities in British Columbia). Previous research corroborated our impression that Canadians of Chinese ancestry are not well represented in our teacher education program (Beynon, Toohey, & Kishor, 1993). We found that although Canadians of Chinese ancestry were well represented in undergraduate enrollments in proportion to their provincial population figures, applications and acceptance to the Simon Fraser University teacher education program of Canadians of Chinese ancestry showed underrepresentation. Current representation in our teacher education program of Canadians of Punjabi-Sikh ancestry was also lower than their university enrollment, which itself was proportional to provincial population figures.² Based on these demographic data, we began an inquiry into the factors that might explain the underrepresentation of these two groups of students.

Our previous research suggested that family members' expectations, parents' wishes, and ethnic group expectations were significant influences on the career choices of visible minority students in British Columbia and were more significant for them than for Canadians of Anglo-European ancestry.³ The questionnaire research showed us that the important factors in choosing a career related primarily to personal, familial, or "cultural" matters and not to externally imposed barriers to career selection. Although students made occasional reference in the questionnaires to issues of discrimination or institutional barriers to teacher education, this format did not appear to elicit many of these sorts of disclosures. The present study was designed to find out more about how a group of Canadian university students of Punjabi and Chinese ancestry saw these factors.

Background

Canadian social policy (Abella, 1984) commonly posits similarities between visible minorities, women, individuals with disabilities, and First Nations as groups in Canadian society who have experienced discrimination by employers and possibly educational institutions. The report suggests that members of these groups should be targeted for training and employment. Social theorists Carty and Brand (1993), McCarthy (1988, 1990) and Ogbu

(1991), on the other hand, advocate the need to look carefully at the specific circumstances of groups whose members experience discrimination. If discrimination is experienced differently by different groups (or by individuals within groups), then efforts to disrupt discrimination may well need to be various. If the reasons that students are not entering teaching careers vary from one ethnic group to another, for example, it would be important to find out what their different reasons are.

McCarthy (1990) makes the point that liberal discourse about race and education overrelies on "culture" and values when explaining inequity, thereby, "abandoning the crucial issues of structural inequality and differential power relations" (p. 56). At the same time, Neo-Marxist approaches, he argues, overrely on economic explanations and imply that individuals make few if any choices or have little agency. A purely cultural view would give the responsibility for change largely to the individual and cultural group. A purely structural view would see the responsibility for change placed largely on economic or other institutions. For example, sociological research on relationships between social class and career choice indicate that "educational plans and vocational decisions are largely determined by socio-economic background" (Li, 1988b, p. 75).

For more than two decades paradigms for multicultural and antiracist education in Canada and the United States have designated the school as an institution with a central role in dismantling inequity. This paradigm casts individual parents, teachers, and students as active and empowered in bringing about change (Banks, 1991; Cummins, 1986; Pang, 1991; Ruiz, 1991; Sleeter, 1991). Specific agendas, theoretical and practical, for transformation of teacher education are also fundamental to restructuring schools, and in turn society at large (Baptiste, Baptiste, & Gollnick, 1980; Bennett, 1988; Beynon & Warsh, 1993; Gollnick, Osayande, & Levy, 1980; Grant, 1983; Lynch, 1986; Martin, 1991).

McCarthy (1988) would have us consider both the personally articulated, "the symbolic, signifying, and language dimensions of social interactions" (p. 246), and the structural aspects, as one considers action around amelioration of racial inequality. The emphasis on personal articulations is consistent with the interactionist approach to social research that focuses on the (diverse) descriptions, explanations, and meanings given by the individual participants. In this view, "to understand people's behavior we must use an approach that gives us access to the meanings that guide their behavior" (Le Compte & Preissle, 1993, pp. 31-32). As Hammersley and Atkinson (1991, p. 7) point out, "It is these interpretations, continually under revision as events unfold [that] shape their actions." Accordingly, in this study, even though we start out knowing, for example, that parental influence is salient, we do not assume that the same parental expectations will have the same effects on all students, even if they are of the same gender and ethnic background.

Our research considers how members of the two most populous visible minority groups in British Columbia describe their decisions about their occupational futures with special attention to teaching. The views they have of themselves with regard to teaching and the influence they perceive their ethnic group membership to have are investigated. Do these individuals feel that they

have encountered discriminatory barriers to their pursuit of education and career? Are some members of the ethnic group more constrained than others in their occupational choices? In what ways do answers to the above questions reveal patterns by ethnic group, by gender, by class, or by some other grouping, and in what ways are responses unique to the individual? Finally, how might this information help us to apply principles of equity in teacher education? By looking at students' views we hope to offer new grounds for dialogue and action in the debate between cultural versus structural analysis.

We have examined in detail the transcripts of interviews of 22 Canadian students of Chinese ancestry and 12 Canadian students of Punjabi-Sikh ancestry. In drawing our data from individual interviews we are aware of critiques of this longstanding research tool that caution about the dangers of generalizing information from individual interviews and transforming it into assertions about classes or groups of individuals in society (Clifford, 1988; Van Maanen, 1988). We are reminded as well that interviewing involves the subjectivities and agencies of interviewer as well as interviewee and that the "answers" of an interview are highly complex and partial (Hale, 1991; Stacey, 1991). Recognizing this complexity and seeing this research primarily as a way to speak to ourselves about how the university might address problems that it is only gradually coming to recognize, we have limited our conclusions, seen them as tentative, and make few recommendations for others. By working with relatively small numbers of participants in lengthy interviews, we also believe it has been possible to be attentive both to expressions of difference and to common themes.

Clearly Canadians of Chinese ancestry have come from many different places at many different times in the past, and the personal, familial, and "ethnic" group experiences of individuals classified as "Chinese Canadian" vary widely.⁴ Differences between the experiences of this diverse group and those of Canadians of Punjabi-Sikh ancestry, of course, are wide as well.⁵ Taking into account all these differences is clearly impossible within the constraints of the present report. Rather, this study focuses on perceptions of teaching as a career by some members of these groups.

Methods and Procedures

The interviews that form the data base for this research were conducted with individuals who indicated their willingness to volunteer on a questionnaire completed by a total of 751 university students at the University of British Columbia and Simon Fraser University (Beynon, Toohey, & Kishor, 1993). The interview questions were designed as follow-up to the topics raised in the questionnaire, allowing opportunities for the students to respond more fully and to indicate the nature of the connections, if any, among their responses regarding family influence, institutional barriers, and career choices.⁶

Interviews were 45 to 60 minutes long. The interviewers were four Canadian women of South Asian, Native Hawaiian, and Anglo-European ancestry. Their experience as interviewers ranged from one to 16 years. After interviews were completed, interviewers discussed the work. Minority interviewers expressed less (retrospective) discomfort in interviewing minority students about discrimination than did interviewers of Anglo-European background. Similarly, they felt visible minority students displayed little unease in speaking about

discrimination with minority interviewers. Interviewers of Anglo-European ancestry were more uneasy asking, and they felt students were more uneasy telling them about discriminatory events. Nevertheless, minority students did report discriminatory events to nonminority interviewers.

Tapes were summarized and transcribed by four women of Anglo-European ancestry (different from the interviewers) with considerable experience in psychology, education, writing, and communications. Each transcriber worked with material from each interviewer and each designated group. They used the format of the interview as their guide to produce a four- to five-page overview with selected verbatim quotations to introduce key points. After reading the transcripts, the authors (two women of Anglo-European ancestry) returned to the tapes to make additional verbatim transcriptions. As a final step in verifying the accuracy and acceptability of the transcripts we sent each to the respective interviewee and asked him or her to correct any inaccuracies. Modifications were minor, as they tended to embellish rather than change the points made. In one instance a respondent indicated that the transcript did not accurately reflect her meanings, and that respondent's interview was not used in this analysis.

We grouped the data first by ethnicity and then by gender. Two comments by students of Chinese ancestry indicate the general awareness that gender was important in career selection:

Traditionally if a girl ends up being a teacher it's great; it's considered one of the best career choices for them, but if a boy ends up teaching that is not so good. A female teacher is prestigious but for a male it is not. (Joshua)

My mom doesn't really care whether I go into anything as long as I get a degree from university. My dad, he actually said, just go get married, who cares about education ... you'll eventually end up getting married ... I have three brothers who are quite good at school so they [her parents] don't really care about me. (Sandra)

Burt, of Punjabi-Sikh ancestry, comments on differences between expectations for boys and girls in his community:

Things are not equal between the sexes in my cultural group and many parents in the Sikh community do not want their daughters to attend university until they are married—if at all.

The sorting by gender within ethnic group is supported by a growing feminist literature that illustrates how minority women's experiences in family, education, and career development diverge in important ways from the experiences of their male counterparts (Bannerji, 1993; Carty & Brand, 1993; Ng, 1993). The report of the students' interviews below makes apparent, we believe, differences in the perceptions, experiences, and aspirations of women and men in these groups.

Our objective in reviewing the interview data was to ascertain both common and diverse themes that emerged from the individual narratives. Although we structure the themes here in skeletal form, nuances emerge in the students' narratives around these issues. The selection of quotations illustrates the salience of the patterns as well as the uniqueness of the individuals whose

narratives revealed these patterns to us. All names have been changed to maintain confidentiality.

We have a great deal of data and finding ways to report these data respecting the particularity of individual's accounts but also considering the length constraints of an academic journal article has been challenging. We make brief general comments about aspects of the information received from students that is relatively straightforward and not necessarily illuminated by the "very words" of the respondents. The body of the data report contains excerpts from the interviews, some of which are particularly apt illustrations of general points and some of which eloquently convey unique perspectives.

Data Analysis

In analyzing our interview data for themes, three factors appeared most salient: although parental influence is one major reason that these students do not choose teaching as a career, we also found that lack of confidence in their own English language skills and experiences of prejudice in the schools are also key influences. A key issue for most Canadian students of Chinese ancestry is parental influence in career choice. Several also identified lack of English proficiency as a barrier to pursuing teaching as a career. Few mentioned racism and discrimination. In contrast, among Canadian students of Punjabi-Sikh ancestry, parental influence and language issues are not identified as impediments to choosing teaching as a career, but racism and discrimination in the schools are. None of the students mentioned that they perceived discrimination in teacher education programs, but this issue was discussed in our earlier publication.

There is no well-developed body of Canadian literature on the area we are studying, but the themes that emerged in our data echo themes in a clearly focused United States body of literature on school achievement for Americans of Chinese and Punjabi ancestry. These studies (as does our work) document that students of Chinese ancestry enter teacher education less frequently than students from other ethnic groups (Peng, 1985; Vetter & Bapco, 1984). The common themes that emerge are the powerful influence of parental expectations on school achievement and career selection, the relationship of language skills to school achievement and career selection, and the reality of prejudice in the schools (Chinn & Wong, 1992; Hsia, 1988; Pang, 1990; Vetter & Bapco, 1984). These themes are also relevant for students of Punjabi ancestry (Gibson and Bhachu, 1991; Singh Ghuman, 1993). A brief discussion of that literature follows the report on our data.

General Comments

Age

The participants as a whole were relatively young. The ages of the men of Chinese ancestry ranged from 19 to 25. Women's ages ranged from 19 to 23. In the Punjabi-Sikh case, ages for four men ranged from 20 to 22; one was 33. The women's ages were from 20 to 22.

Class

The class background (defined here by parental occupation) of the students ranged from laborer to professional.⁷ In the case of the men and women of Chinese ancestry, parental occupations included cook, baker, seamstress,

homemaker, accountant, teacher, pharmacist, doctor, and entrepreneur. In the case of the men and women of Punjabi-Sikh ancestry parental occupations included millworker, farm laborer, machinist, homemaker, bus driver, teacher, social worker, and engineer.

Financial inaccessibility of postsecondary education is cited often in sociological research as a critical factor preventing minority youth from attaining high socioeconomic status through the pursuit of managerial or professional careers (Breton, 1970; Buttrick, 1977; Gilbert & McRoberts, 1977; Li, 1988b). Quantitative data from our previous study indicated that this is not necessarily the case for students from the two groups studied, who come from a wide range of socioeconomic backgrounds and are well represented in university (Beynon et al., 1993).⁸ Class background of the students may well have an effect on both the sorts of careers they select and the reasons they find compelling for choosing those careers. We report parental occupations in the sections on parental influence, but the data from this study did not reveal any clear relationships between parental occupation and perceptions of the desirability of teaching as a career.⁹

Countries of Origin

Related to issues of class are the diverse economic histories of families in their countries of origin, which may differentially influence career choice. In contrast to earlier policies emphasizing recruitment of labor, Canadian immigration, particularly since 1979, has emphasized recruitment of business immigrants with a large amount of capital to invest in Canada (Li, 1988a, p. 125). For Chinese immigration this has meant an increase in business immigrants from Hong Kong. However, of the students we interviewed there were only two instances (of students of Chinese ancestry) where parents might be identified as post-1979 business immigrants.¹⁰

The career choices the students had made are reported in Table 1.

Canadian Women of Chinese Ancestry

Only one of the 12 women in this sample was planning on teaching as a career. We begin with a profile of her experiences. This is followed by an analysis of the other 11 women who framed their reasons primarily in terms of parental influence and English language proficiency. In this group of women only two reported experiences of prejudice and discrimination in their own schooling.

Table 1
Career Choices of University Students
of Chinese (cc) and Punjabi (cp) Ancestry

	Law	Business	Science	Health	Engineering	Public Relations	Arts	Education	Police Work	n
cc (female)		4		4			3	1		12
cc (male)		4	3	1	2					10
cp (female)	1	2	1			1	1		1	7
cp (male)	1	2	1					1		5

Lydia, who is enrolled in a teacher education program, is 22 years old and immigrated as an infant from Hong Kong. She told us about when she first thought about becoming a teacher.

In grade 1 my teacher was reading us a story and I liked what she was doing so much from that moment all I thought about was teaching.

In spite of difficulties encountered in school and ambiguities in her parents' expectations she was still committed to pursuing a career in teaching.

In high school it was very difficult for me and a lot of teachers told me that I wouldn't get very far, that I wouldn't get past high school. I resented that. My grade 9 and grade 10 math and French teachers told me not to go on in school. If I teach I want to be more effective; I want to be more nurturing. I want to let students know that if they aren't doing well right now it's OK, there's always a second chance. You can always improve.

[My parents] didn't think I could do well enough at university, although they think teaching is an honorable job and are fairly satisfied with my choice [to be a teacher], but they think I am going to be beaten up [working in a school]. (Lydia, parents' occupations: grocers)

Parental Influence

The women cited strong and pervasive parental influence and expectations, and in general parents seemed more willing to accept teaching as a career choice for females than for males. They distinguished two realms in which parental influence and expectations operated: university attendance and career selection. Ten of the 12 women's narratives begin with strong statements that as long as they attended university, their parents "didn't mind" what career they followed. But as their narratives unfolded it became apparent that their parents favored careers in business that were deemed practical and financially secure. Emphasis on occupations deemed prestigious did not appear to be as pronounced as it was in the case of the men. However, there were comments indicating that for women university attendance in and of itself was considered prestigious and desirable with 10 of the 12 women citing explicit parental expectations about university attendance. Typical was:

Ever since I was a kid it was assumed I would go to University. The Singapore mentality on the whole is, your job is to study, to get good grades and that is that. (Lauren, father: entrepreneur; mother: homemaker)

I just knew that I had to go to University; it wasn't a question of whether I didn't or not. It wasn't anything I could pick. I couldn't say that I wanted to take a year off and go to work. It was assumed I would go to University. (Alice, father: pharmacist; mother: homemaker)

My parents would rather have me get a degree than just a diploma or a certificate. When they talk to their friends and relatives and say "Oh, my daughter goes through university" instead of just a college or an institute, it just sounds better. (Sandra, father: retired; mother: homemaker)

Eight of the 12 women noted parental approval (in varying degrees) of teaching as a career for them, occasionally referring to someone in their family who already was a teacher. A sample of their comments:

They think it's a good occupation and they wouldn't mind at all. (Corinne, father: gasfitter; mother: property manager)

My father gave me his blessing with teaching because my mom is a teacher. (Lauren, father: entrepreneur; mother: homemaker)

My parents would approve of teaching, just as equally as nursing. My sister is a teacher. (Lucia, father: chef; mother: worker)

My parents would be neutral. It is a secure profession. To my mother's generation teaching was the most natural thing to do for a girl. I'm not too sure about my dad, but my mom wouldn't mind. (Carol, father: translator; mother: office worker)

Two of the 12 women reported their parents would not approve:

My family would approve of teaching only if there is a high demand for teachers in the marketplace. It is not a preferred career and the first thing they examined was the financial aspect. (Laura, father and mother both retired)

My parents would think there was nowhere to go as a teacher. (Veronica, father: machine operator; mother: computer service clerk)

There was greater diversity of career choices among the Canadian women of Chinese ancestry than among the men. Those women whose choices were seen by parents as impractical or not leading to financial security (e.g., French, Asian studies, and film studies), reported conflict with parents or feelings of inadequacy, dissatisfaction, or uneasiness with self. In a few instances both university attendance and career choice for girls were seen by one or both parents as subordinate to the goal of marriage and family.

I had originally said when I was in high school that I wanted to go into medical school and my mom said, "Oh no, girls don't do that. You'll never get a husband that way." (Lucy)

Language and Communication

Several of the young women in this sample mentioned lack of proficiency in English as preventing them from pursuing teaching as a career. Jill, who had attended high school in Canada for one year, felt teaching was a career for people who grew up in Canada:

I have never thought of teaching as a career. Maybe it is because my language [English] is not good.

As two others put it:

I have a problem talking in front of a lot of people, so I think I'd have to overcome that fear first. I get really nervous, usually with superiors or peers my age, but I guess with kids a little younger than I am I wouldn't be afraid but with high school level I would be. (Corrine, age on arrival: 3)

I tend to mumble a lot, especially for words that I know I couldn't pronounce it right and I tend to get nervous ... I don't think I would be a good teacher in Canada because of my language ... I still have quite a heavy accent and I don't think it's good to be a teacher and have such an accent ... in a lot of cases, I couldn't pronounce words correctly, so I didn't choose teaching. (Sandra, age on arrival: 11)

Discrimination

The majority of this group of women of Chinese ancestry reported no experiences of discrimination in schools. Laura, however, was eloquent in supporting her view that “generally school is not a pleasant experience for minorities.”

I got into trouble in preschool, that is where I experienced my first racial remark, the first time I knew what racism was. I knew what the words were, I knew that they were bad, because my brothers and cousins would talk about it. And I met this girl and we were on a slide and she called me a chink and a Chinaman, and she was standing up on the slide and I pushed her off, and she landed flat on her face, and I got into trouble and I have been in and out of it ever since.

As a child you are intimidated. There is this tall white lady telling you that you have to go to ESL, even though you watch Sesame Street and can say your numbers in French.

Josephine remarked:

We were always wondering if teachers were wondering why we [Chinese Canadian students] were here [in French class]. Again you had to be top-notch. [Josephine had earlier explained “You have to be 110% to be accepted in the white world.”] And I couldn’t stand it. I’m here to learn, I’m not here to become you.

Canadian Men of Chinese Ancestry

None of these 10 men is currently preparing for or interested in a career in teaching. Only one respondent could remember ever having wanted to be a teacher. When questioned about why they were not interested in teaching careers, difficulties with language and communication skills were seen as a prominent barrier; school climate and discrimination were also mentioned by some of the respondents. Parental views and influence were identified as the background against which these other factors played themselves out.

Parental Influence

Most of these men acknowledged strong, consistent, and pervasive family influence to choose jobs from professions deemed financially secure and prestigious (such as medicine). If this was not possible, second choices were in the areas of business, computing, and accounting with teaching at the bottom of the list. These expectations are stronger for men than for women. The sons responded to these expectations in a variety of ways, generally finding means in the end to pursue their own choices. However, no one selected teaching.

In the beginning, my parents’ expectations were very high. They wanted me to go into medicine or into business. But it’s a bit difficult. I found getting the grades to go into those fields difficult. But now my parents just sort of let me be, sort of. I can’t get into medical school or the business faculty so right now my parents have taken a really relaxed attitude. They say, “Well do your best and then we’ll see what happens from then.” There used to be a lot more pressure from my parents to do certain things. (Peter, father: retired doctor; mother: homemaker)

Basically, our family has always been in business so there is a little bit of pressure on going into business. (Hayne, father: investor; mother: housewife)

[My parents] did not influence my career choice [Wesley—the only man in the sample going into medicine] and in my teen years I was sort of a rebel so I didn't really listen to my parents, but they did influence me to attend university. (father: baker; mother: laundry worker)

There are pressures from my family to get a professional career— either a doctor or a lawyer. They felt more strongly about this before but now I know what I want so they have less of an influence. [Richard is now pursuing a business degree] ... My mother always thought business was a risky thing versus a doctor ... If you got shipped back to Germany or something, you would always have a career. (father: cook; mother: housekeeper)

As Richard points out, these parents' views of desirable occupations may be importantly formed by considerations of career security (and perhaps portability). The impact of parents' views of desirable occupations was reinforced by the financial resources that they provided to make university education possible. Regardless of parents' occupations or level of education, all but one of the interviewees said finance was not a barrier to higher education. Typical was:

Money is not a factor because in my family an education is sacred and they will lend me the money or even borrow the money for my education ... they had a scholarship fund for me when I was a baby so I have to at least get my first degree. (Wesley, father: baker; mother: laundry worker)

All these men lived with their parents and this was noted by some as an important cost saving factor.

None of them saw their parents promoting teaching as a career of choice. Doug, the one male student in this sample who had indicated an early interest in a teaching career (but later made another choice), said:

When I mentioned to my mother that I might want to become a teacher she questioned how good teaching would be as a career. (father: noted as deceased no occupation given; mother: teaching certificate from Hong Kong, currently a homemaker)

A number of respondents noted that their parents' views of teaching in a Canadian setting ran somewhat counter to the notion that in Chinese culture teaching is considered an honorable occupation.

Well in Chinese circles it *used to be* highly respected to be a teacher. (Wesley)

I always found it ironic that although, according to classic Confucian ideals, a teacher is accorded respect, in real life this hardly ever happens. (Wallace)

Language and Communication

Four of the nine men mentioned they had difficulties in the area of language and communication, which precluded their considering teaching as a career. Three of these four men had grown up in Canada. As Peter who was born in Canada put it:

You have to have good speaking ability. You have to be able to communicate. You have to be able to express your wishes to students or whatever and you have to deal with the students on a daily basis. I don't know if I can do that.

Discrimination

Two of these men had a substantial amount to say about discrimination in educational institutions. Jason attended a multiethnic urban secondary school and noted:

It was very bad between students. You learn to take care of yourself. Name calling is always big. Mostly between Chinese and Caucasians was the major conflict. I think I was pushed around quite a bit because I was Chinese. A scapegoat. Day-to-day things ... Teachers ... accepted it as something that happens. They didn't really do anything about it.

Wallace said that negative experiences ("social problems") in high schools made returning there for work an unattractive option:

Minority students are not interested in teaching because of the social problems that exist in schools, with gangs, drugs, and the large immigrant population. Many minority students did not have good experiences in Canadian high schools and would not wish to go back to that same environment. Many Asian students do not integrate very well and tend to stay in their own cultural groups. Consequently they get bored and turned off and do not get very involved with the social community of the school. They are then unlikely to want to get reinvolved in schools after they graduate.

Some of these men reported no problems at all with discrimination in their schooling. Wallace was equivocal about his experiences:

If I had had anything [discrimination] like that happen to me I would have just said to myself, got to try harder. It may very well be that I did come across that ethnic thing. If anything bad were happening in my life, my response would be try harder, try harder, try it again, overcome, be popular, try to be popular ... get on student council, don't wear plaid shirts and brown pants every day. That was my response to any sort of thing.

He then described an experience in kindergarten where other boys repeatedly stole his lunch for the entire school year.¹¹ He continued:

That was one experience that pushed me to say I'm not going to take this anymore. Whatever the world throws at me I'll take it and I'll do better 'cause I have to 'cause otherwise where's my family going to go? ... That never got solved and bothered me. I decided I'm not going to let this happen to me again. I don't want any of this to happen to me or my family again.

The five men who reported that they personally had not had any difficulties with discrimination said that it was not an issue of any importance to them in making decisions about their careers.

Canadian Women of Punjabi-Sikh Ancestry

One of the seven Canadian women of Punjabi-Sikh ancestry in this sample was preparing for a career as a teacher. When the other six were questioned about their career decisions, parents, especially fathers, were perceived as influential. In general, these women did not appear to evaluate their language skills as currently inadequate to do any kind of job, although some readily acknowledged English language difficulties in the past. Concerns about discrimination were commonly expressed.

The young woman in this group who was at the time of the interview, preparing for a teaching career, was alone in citing high school teachers as encouraging in terms of education generally and interested in her career choice. She described experiences with two teachers where she had been encouraged not to quit school and to consider either law or teaching as possible careers. In the follow-up mailing this woman indicated that although she had been accepted to a teacher education program, she had finally enrolled in law school.

By doing volunteer work with teens I found I didn't have as much patience or interest as I thought I should have had ... I was also discouraged quite a bit by old high school teachers who felt that teaching was no longer teaching but just babysitting. I must admit this had an impact on me. I began to wonder if I had what it took to avoid adopting such a pessimistic view of teaching. (Pindy)

Parental Influence

All these young women indicated that their parents' opinions were of concern to them. Six of the seven women cited their fathers as the primary influence on their career decisions. The seventh woman's father is deceased. One woman put it this way:

I get a lot of my ideals from my dad, you know, this idea that I have a responsibility toward community and to put something back and to help change people's lives. I am very politically committed and that comes directly from my father. (Rajinder, father: social worker; mother: home manager)

Another woman studying medicine said:

My dad's encouragement was always positive ... I always wanted to make my dad proud of me because I am really close to him. (Surjeet, father: mill worker; mother: homemaker)

These women singled out their fathers as influential in terms of their career choices, often remarking that their mothers were more concerned with their "happiness." As one put it:

My father has always been the one to encourage me to do well. Always since I was 6 or 7 years old. My mother was always happy if I did well and it was important, and if I didn't I would hear from her, but I talked to my dad when it came to school. The general idea of having done well and getting a good job in the end was important. There was definitely a difference. My mom would want me to be happy and my dad would want me to be successful. I was talking to my cousin and she said almost the same thing. I don't know if it is a general female versus male stuff. (Jathinder, father: MA degree in math, works as a bus driver; mother: has a teaching certificate, works as a secretary)

Three of the seven women indicated that their parents would find teaching an acceptable career for them. In one case the daughter indicated she thought they would be quite pleased with this choice and related that both parents had been teachers in India but were now working respectively as a secretary and a bus driver. One indicated that her parents would like it better than police work (which she is pursuing) because "they would think it would be safer." The third said that teaching is highly respected, but that "the money wasn't that good."

Language and Communication

Three of the women in this sample mentioned English language difficulties when they first came to Canada, or when they started school. Kuldip (now 22), who came to Canada at age 14, reported that she had initially had difficulty acquiring her first (part-time) jobs due to language barriers. After some time, however, she felt no further difficulties in this area and did not mention any difficulties with English as being involved at all in her career choices.

Surjeet, born in Canada with Punjabi as a first language, describes difficulties when she first started school:

I used to cry and cry. I was very shy. I couldn't speak English and I went to this humongous school. I was always nervous about my English. There were new immigrants and I wouldn't talk to them because I felt they just got here and it wasn't cool.

Surjeet is now enrolled in medicine, and English language proficiency was not mentioned as a factor in her career choice.

Sarita (now 20) immigrated to Canada as an infant and described her first languages as Punjabi and English. Her mention of language is ambiguous:

After grade 12 I was very scared as to where I was going to go. And I didn't have the language so I had to go to college.

After deciding that the college courses were not of interest, Sarita enrolled in criminology at university and mentions no further difficulties with language.

Discrimination

In the questionnaires only one woman wrote about experiences with racial discrimination, but in the interviews six of the seven women spoke about these matters. These young women reported that the most overt hostility they encountered was from other students rather than teachers. However, teachers rarely intervened when they saw racist events between students. The following excerpts from interview transcripts show that they perceived discrimination occurring across the grade levels in a variety of ways and that coping strategies were various as well.

Students were very cruel ... The teachers were OK ... They were supportive. They went out of their way ... Things have changed. When they came here they [Sikh parents] didn't have the facilities. They all came from villages. They had to assimilate, they had no choice. They had to bring up their children this way. My mother felt like she had to wear a miniskirt, otherwise people would harass her walking down the street in a sari. It is a double whammy if you are adolescent and a minority. (Surjeet)

I remember one [high school] teacher saying, "You can't see her blush because she is so dark." And you don't know how to take it because this person is a teacher and they're laughing and it's a joke ... [But] most teachers were fairly good. (Rajinder)

Rajinder also expressed interest in education, but not in a public school setting because she felt intolerant of the racism of youngsters:

The kids in the education system at the moment are very obnoxious. Maybe not all of them and in different schools it might vary, but I just had this experience working with a group of kids for a few weeks and they were so racist, I couldn't believe it ... Actually there was one person in the group who was also Indian and they would make comments to him, and then they would make sort of oblique comments to me about the color brown. It was just obvious from the tone of their voice that they were being racist and I couldn't deal with it at that point and I really didn't want to.

Jathinder spoke about discrimination in this way:

There were incidents in grade 3 but it never affected me in any way although it was to people of Indian origin. It didn't bother me. It never happened in front of the teacher so you don't go running for help. I just thought they were really silly and immature. Nothing serious ... One thing I remember from high school was from a teacher we all hated, everybody. Generally people consider me to be a shy person or quiet, so I don't know whether I answer something loud enough. This teacher said to me, "You don't have to be so quiet because you are Indian. Maybe it is expected of you in your culture but ..." I don't think of myself that way. A lot of people think Indian women are submissive and I am not like that and neither is my mother. I don't know if he was trying to be rude, we all thought he was a jerk anyway. In university in commerce ... you have to take on Caucasian values to succeed ... It is nothing the teacher says or anything but institutionalized racism. It may also have to do with gender.

Harjeet shows her strategy for avoiding racism:

I experienced less racism in grades 11 and 12 because I didn't hang around with other East Indians. When I was in grade 7, we had a girl who came from India and they asked me to help her out, my being East Indian and everything. That's what they have ESL teachers for ... I don't think students should have to bear the burden of teaching another student a new language ... Kids would say or imply that I should stay with "your East Indian friend."

Rajinder spoke of her impatience with a university class in which she felt the professor and other students were avoiding discussion of clearly evident racism:

They don't talk about realities in a lot of arts classes. I just had an English class last semester and we were discussing some very political texts, and I watched and no one said the word racism for three weeks. That really frustrated me because obviously that's what these texts were about. Toni Morrison, Alice Walker, that's what these people talk about. (Rajinder)

Canadian Men of Punjabi-Sikh Ancestry

One of the five men interviewed here is preparing for a career in education. In choosing a career Rick's first motivation was to earn a lot of money, so he entered the Faculty of Business, but he soon realized based on his experiences teaching soccer and assisting in a day care center that he valued other things, such as working with children. His parents were happy with his choice.

The men who were not interested in teaching as careers had made rather detailed plans for their futures, involving short- and long-term goals. Although questioned specifically, only one of these men gave a clear reason for not choosing education, which was his father's discouragement. These men did not indicate that they felt any personal inadequacies had precluded their choosing

teaching as a career: only one, for example, mentioned English language proficiency as a personal problem, but one he had solved. The sense one has from reading transcripts is that these young men have simply made other choices.

Parental Influences

Four of these five men, like the women, express the importance of their fathers' wishes in their career choices.¹²

I want to be something that would be respectable in my dad's eyes. My dad wants me to become an accountant because he thinks that is a good job to have. He just wants me to graduate with a good degree so he can say "my son is graduating with so and so." There is a bit of pressure but not a lot. My mother doesn't think about it too much. She just wants me to graduate. My dad really talks to a lot of people and he knows my personality really well ... So I go to my dad for a lot of advice. (Rick, father: assistant engineer; mother: chef's helper)

Language and Communication

None of these men cited language difficulties as being implicated in their career decisions.

Discrimination

With regard to issues of prejudice and discrimination, these men had a variety of perceptions.

My whole life I have hung around with English people. I find if I have a white friend I can trust him more, because it tends to be that a lot of East Indian people if they don't like you they will start rumors about you. There is a lot of jealousy. I had no problems in my school because I knew a lot of English people, so I was always treated like one of the guys. In elementary there was not a great understanding of religion but it didn't matter. Some kids you get bugged by just because you are different. I was glad I didn't have to wear a turban ... I was very lucky. I have great parents. They're very westernized. My parents can talk to my friends. (Rick)

Anoop said about discrimination:

I just think there is no communication there [if minority students are treated unfairly in school]. If I have a problem, I go up and tell the teacher. I say, "Listen, why is this the case? Why am I doing poorly?" and usually the teachers are very helpful in that sense.

He did comment on racial differences and teaching, however:

I've always associated teaching as a white profession.

Perry (whose father is a high school teacher) dealt with the question of personal experiences with discrimination by talking about what he thought teachers might do to alleviate such difficulties:

I think that it should be a teacher's responsibility, especially if they are teaching in a multicultural district, to understand the cultures a little bit more ... it goes beyond the significance of why they wear a certain type of clothing or their food and dancing; religion is a really touchy subject, but it is such a fundamental subject that I believe that kids should be exposed to why people's beliefs are the way they are, it just leads to a bit more understanding, teachers have to play a bigger role. I wouldn't limit that to ethnic groups, I would say it's just social issues. Making sure that teachers treat minority students well while they are in

grade school is an important first step in encouraging minority students to become teachers.

Discussion

Parental Influence

These students' reports as a whole indicate that their parents' play two important roles in their career choices: one is financial; the other relates to views of the desirability of particular careers. Although some of the students had to work or live at home in order to attend university, almost all indicated that their parents' support for postsecondary education was such a high priority that they would find the money for their children's education regardless of their socioeconomic circumstances. For these groups the relationship of social class and educational attainment is different from that found in earlier large-scale Canadian surveys showing that students from families of higher socioeconomic status are much more likely to attend university than their peers from lower socioeconomic backgrounds (Anisef & Okihiro, 1982; Porter, 1965; Porter, Porter, & Blishen, 1979).

Parents also play an important role in selection of specific careers. For Canadians of Punjabi descent it appears that parents' views, especially their fathers', are important in decision making. The greater influence of fathers in determining educational goals and choosing a career is also identified by Wakil, Siddique, and Wakil (1981) in their study of Indian and Pakistani families in a western Canadian city.¹³ The students we interviewed of Punjabi-Sikh ancestry indicated that their parents would have no strong objections to teaching as a career choice. In the cases of the one female and one male who had chosen it, there was parental support. The males of Chinese ancestry reported their parents would not be supportive of a decision to enter teacher education. In the case of the females of Chinese ancestry, eight of the 12 women reported parental approval of teaching as a career choice for them, but only one of these young women has chosen this path. The one woman of Chinese ancestry who was preparing for a teaching career reported her parents were *not* supportive of this aspiration (nor had her teachers been encouraging in this regard). In general, then, the parents of students of Chinese ancestry are not supportive of teaching as a career for their sons, but are less unified in their opposition to this career for their daughters. Nevertheless, these young women are generally not choosing teaching as a career for themselves, sometimes because, they report, of inadequate English language skills. This matter of language skills is discussed below.

Language and Communication

The careers identified in our study as desirable to males and females of Chinese ancestry are also those where quantitative technical skills take priority over linguistic or communication skills. Some of these students are native-born, whereas others were older on arrival in Canada. The predominant number of students of both Chinese and Punjabi-Sikh ancestry grew up in families where the English language played a minor role. Despite many years of English medium schooling, several cite English language difficulties as reasons for not choosing teaching (or other careers for which verbal skills in English are necessary). These difficulties with language are mentioned often in American

literature concerning students of Chinese ancestry. Pang (1990) notes with regard to these individuals, "They not only feel the inability to do well but reveal a fear of writing and speaking" (p. 58). Hsia (1988) states that the contrast between Asian Americans' achievement in quantitative fields and their avoidance of the difficulties in the fields that demand well-developed verbal skills is stark among recent immigrants and is still noticeable, even after several generations among the native born (Chinn & Wong, 1992). Chinn and Wong argue that overlooking the relatively weak verbal skills of Asian students and making no educational provisions for improving those skills is unfortunate because

Early development of these skills can enhance social interaction with non-Asian peers and reduce the avoidance of courses out of the quantitative areas. The development of verbal skills, coupled with the acquisition of social skills, can enhance the likelihood of Asians branching out from their traditional areas of study and thus help in their recruitment into teacher education programs. (p. 131)

Unlike the Canadian students of Chinese ancestry, none of the students of Punjabi ancestry mentioned current difficulties with English as responsible for their decisions not to pursue teaching careers. Some acknowledged previous difficulties, but all appeared to regard these as difficulties they had overcome that were no longer relevant.

Discrimination

Some of the Canadian students of Chinese ancestry spoke about experiencing discrimination in their own educational experiences, and two implied that racial, as well as other tensions in the Canadian public school system made them unattractive work places for minority students. Although some clearly and strongly articulated difficulties they had had as students in schools, others had little to say in this regard and said it was not an issue of any importance to them in career decision making. The American literature on students of Asian ancestry stresses the need for schools to be sensitive to their emotional and social needs (Pang, 1990). The point is made that students who feel powerless to deal with prejudicial situations may withdraw from the school community, thus diminishing the likelihood of pursuing careers in teaching.

The Canadian students of Punjabi descent were also diverse in their responses to questions about prejudice and discrimination. Five of the six women interviewed described incidents of racism in their own schooling, and one indicated that the racism of students in the school system was part of why she was not personally interested in a teaching career. The Canadian Punjabi men had less to say about these matters. They reported they had personally dealt with racial discrimination by ignoring it or by avoiding identification with their own racial group. These men displayed little interest in discussing these matters.¹⁴

Interpretation and Recommendations

We began this investigation reminded by McCarthy (1988, 1990) that diversity in the experiences within and between minority groups needs to be recognized and investigated. In looking in detail at the experiences of individuals from the two most populous groups in British Columbia (and looking at them by

gender), we were expecting diversity. This expectation was not unfounded. We were continually struck in reading the interviews and listening to the tapes by the many ways individuals made decisions and managed problems as they set about their task of getting the education they needed to make a living. At the same time, we also noticed commonalities in the experiences and perceptions of members within each of the four groups described. How does the information we have gathered and our analysis of it affect our perceptions of equity policies and barriers to the representation of Canadians of Chinese and Punjabi ancestry in the field of teacher education? How can we improve? Our analysis has identified three main dimensions: parental influence, language and communication, and discrimination. We discuss each in turn below.

Parental Influence

For both groups studied here, reaching out to parents is important, but it appears that what parents need to know will vary somewhat between the two groups. When the men of Chinese ancestry were asked what could be done to encourage more minority students to go into teaching, suggestions focused on prestige and finance:

You have to influence the parents to change their ideas and also to have better financial incentives. (Peter)

Sell the notion of giving it a prestigious status. (Wallace)

Enhance the perception of teaching in the eyes of Chinese parents. (Doug)

Although faculties of education clearly will not in the foreseeable future affect the remuneration offered to teachers, it is apparent that if recruitment into faculties of education from these minority communities is desired, communication with parents will be important. Outreach to minority community organizations and through the ethnic media are avenues that have not yet been tried but that might yield positive results. Media and public information targeted at communities should be available not only in English, but in Cantonese, Mandarin, and Punjabi and should include perspectives on these careers not only from students, but from parents and respected community leaders including those who have successfully pursued careers in teaching. We are aware that many outreach efforts to the community at large focus primarily on the personal satisfaction and fulfillment associated with a career in teaching. In view of differential concerns between and among groups identified in this research, it would be important to be specific about a wider range of occupational concerns including socioeconomic status (e.g., financial security and prestige) as well as personal esteem, satisfaction, fulfillment, safety, and service to the community.

As our research has indicated, there is fairly strong resistance to teaching as an appropriate career choice for men of Chinese ancestry, and so attention will need to be paid to this gender issue in recruitment activities. In addition, because there seems less resistance to women going into teaching careers, it may be that recruitment of women of Chinese ancestry will be less problematic than in the case of men.

In addition to the above considerations statements by mainstream and minority educators should deal forthrightly with issues of language acquisition and discrimination in the schools.

English language proficiency may be regarded (as many of the Canadian students of Chinese ancestry see it) as a personal problem or a “cultural” pattern. Students of both Punjabi and Chinese ancestry identified English language acquisition as problematic, with only the Chinese seeing this as persistent. As indicated by the data presented earlier, difficulties were experienced by those who emigrated in their teenage years, but also by several who were born in Canada.¹⁵ Although this research does not inquire into the reasons for the differences in persistence of problems between the two groups, it would seem that linguistic as well as social factors may account for the variance.¹⁶

It is important also to remember that career decisions have not only to do with differential socialization, talents or interests, but also with opportunities and institutional arrangements. Seeing Canadian males and females of Chinese ancestry as being “naturally” and “culturally” concerned with science and business, and absent from arenas where work is importantly conducted through oral communication, for example, may ignore structural arrangements that make it difficult for these young people to enter these professions. What they identify as linguistic barriers might encompass what they also perceive as social barriers. This research has clearly shown that a number of individuals from at least this group of Canadian students of Chinese ancestry self-evaluate themselves as relatively unskilled verbally and unable to take on verbally oriented careers. Are the students right: are their verbal skills really less developed than would be necessary for jobs like teaching? Why would a group of students who have lived in Canada for many years, attended English-medium schools, and have enough intellectual ability to attend university not have developed correspondingly in English language proficiency? Have individuals from this particular background been excluded from certain kinds of conversations? Have they been somehow persuaded that their English language abilities are inadequate? Questions like this suggest that what might appear to be personal and individual inadequacies may be constructed impediments to the participation of individuals from particular groups in particular jobs. This structural analysis is supported by Li (1988a) who identifies relative absence of Canadians of Chinese ancestry in jobs requiring social interaction as a strategy for avoiding disadvantageous competition.

Distinguishing between cultural and structural sources of lack of participation may be an intellectual nicety of little interest to the students concerned. If high schools wish to increase the options in career choices of all students, it may be that English language courses need to be available more widely than heretofore. More emphasis on oral communication through drama and role-play would be beneficial in preparing for a teaching career. If faculties of education are to encourage more minority students to enter and complete teacher education, it may be that those faculties need to offer ways for some of those students to improve their English language skills or to increase their confidence in speaking in certain arenas. As we prepare anglophone teachers to work in French immersion settings, we offer courses explicitly designed to improve candidates’ French language proficiency. If universities are to prepare

minority students to become teachers, students may need opportunities to improve their English language skills so that this job requirement does not preclude their participation.

Discrimination

Clearly individuals in each of the four groups have experienced discrimination differently, and have learned to regard and deal with it differently. The Punjabi women students appeared to be most acutely aware of it. (Our research was not focused on reasons for variation in articulations of discrimination). They may relate to real differences in discrimination experienced by the two groups or to different levels of awareness. It is difficult objectively to relate and compare the severity of discrimination, but at least in recent years the Canadians of Punjabi-Sikh ancestry have been the focus of much media attention in this area.¹⁷ Many of the students have articulated their perceptions of secondary school incidents of discrimination on the basis of race or ethnic identification. In some cases, students said that they perceived schools to be unwelcoming to “people like them,” and that it was because of this they had no wish to work in such places. None of the students we interviewed had participated in multicultural leadership camps that are available (on a limited basis) in their schools, and none spoke of multicultural programs or curriculum. The one student who mentioned the positive influence of a minority (teacher) role model was referring to his father.¹⁸

Not all the students we interviewed experienced overt instances of personal discrimination. In our view all experienced implicit institutional discrimination in schools that took little notice of their unique positions. The school lives of these students appeared to be untouched by the variety of multicultural and antiracist programs and curricula that have a “social change mission,” and “link power and empowerment with race, social class, and gender issues” (Sleeter, 1991, p. 2). They could not refer to any examples in their experience of the four key areas for social transformation of schooling and empowerment of students identified by Cummins (1986).¹⁹ In spite of this, they were successful in the conventional norms of educational institutions. But we see their reluctance to pursue teaching as a measure of the ways these institutions implicitly exclude/deny and discriminate against them. Clearly, if minority students are not to be discouraged from working in schools as teachers, the climate in public schools for minority students needs to change.

In spite of the apparent absence of contact with multicultural/antiracist programs, the students we interviewed had many ideas about what teachers could do to improve the climate in high schools. For example,

Make others aware that there are other cultures here ... I guess they [other people] find that some of the things we do are weird or different and they don't understand why. (Veronica)

If students feel differences ... they need to feel safe enough to talk ... but not to have the issue blown up into something major, but to have it dealt with ... When teachers try to be multicultural, it can make people proud or it can make them more self-conscious. You don't want to be made to feel self-conscious and different. But on the other hand you don't want to be ignored. It is very sensitive, it's very hard not to do it without generalizing and flattening, it takes a lot of commitment from the teacher. (Rajinder)

Because much of the responsibility for school change falls to teachers, the absence of comprehensive programming in multicultural/antiracist teacher education is troubling and problematic (Beynon & Warsh, 1993; McCarthy, 1990; Sleeter, 1992). This area is discussed in the concluding section.

Conclusion

We agree with Fleras and Elliott (1992) and Henry and Jain (1991) that employment equity legislation directed at the hiring process is important. However, this legislation (and what appears to be the inevitable resistance to it) must not divert attention from the necessity of initiating programs for educational/training equity that are attentive to diversity among and within groups and consequently provide students with appropriate preparatory skills and experience.²⁰

In this research we have been guided by the interactionist emphasis on ascertaining the meanings that guide individual behavior and McCarthy's (1988, 1990) emphasis on diversity between and within ethnic groups. These have been useful in providing insights into the ways social institutions and cultural influences impinge on individual perceptions and actions regarding the choice of teaching as a career. We have seen that at least two prominent visible minority groups have some commonalities in regard to parental influence, but also that they are different in important ways (e.g., their experiences of discrimination and their self-assessment of their skills in English language and communication.) As policy analysts and implementers we see the knowledge of the group and the search for commonalities serving as a heuristic in identifying where and why inequities may exist for visible minorities vis-à-vis mainstream Canadian institutions. As educators we are acutely aware of the need to view a group in interactionist terms through the descriptions, meanings, and explanations provided by individual students. Knowledge of the individuals will help both in preventing the group perspective from becoming a stereotype and in designing educational initiatives that will respect the needs and concerns of the individual students.

Thus as policy analysts, implementers, and educators we need to become adept at using analytical bifocals so that we can shift easily from a group to an individual perspective. With this general approach in mind we turned to specific recommendations for improving the representation of minorities in teacher education. Even with this small-scale study we see individuals who, in spite of institutional barriers such as inadequate language training and garden variety racism, *do* wish to pursue training or employment in teaching. The support, encouragement, and outreach of mainstream educational organizations such as the teachers' federations, universities, and ministries of education should play a stronger role in assisting them, and other minorities as well.²¹

Our research also addresses some practical issues related to making closer links between the world of school and the world of work. Several provincial governments and ministries of education in Canada are embarking on renewed efforts to have students develop work-related skills prior to entering postsecondary education. In British Columbia the Ministry of Skills Education and Training (formerly Advanced Education) and the Ministry of Education have joined forces in launching the Skills Now initiative, which builds career and personal planning and work experience into the curriculum in increasing

percentages from grade 4 through to graduation. These initiatives acknowledge the necessity of a partnership with parents, but are otherwise vague as to what this should entail. Our research indicates that, at least as far as teacher education is concerned, these partnerships might need to be as varied as the parent groups to whom they are reaching.

Our research also supports the intent of both ministries to provide youth with positive school experiences, including mentorships and internships that support an informed process of career selection. New efforts in this area need to take special care that students from all groups are reached and that this may require unique approaches including collaboration with community-based organizations and youth groups that generally operate outside school boundaries.²² Such efforts are necessary in order to provide young people convincing evidence that they are valued both as students and as future members of the teaching profession.

Although we view these government initiatives as potentially useful and supportive of minority students, we also fear that they may have little effect unless pervasive concerns about school climate are addressed. As long as this climate does not actively support students through multicultural and antiracist as well as other kinds of programming; as long as teachers are not routinely prepared by universities and professional associations to take an active role in transforming their classrooms, we doubt that new government programs will reach all students and provide them with equal opportunities.

Notes

1. Concern about equity of access and equality of representation in various fields of employment has been evident in Canadian society at least since the Royal Commission Report on equality in employment (Abella, 1984) and the subsequent passage of the federal Employment Equity Act (Government of Canada, 1986). Universities and other federally funded employers, for example, are required to review the composition of their work force with a view to determine whether it is representative of First Nations, visible minorities, women, and persons with disabilities; if it is not representative these institutions are required to develop plans to address this imbalance in future hiring.
2. These estimates of representation are based on averages of 10 semesters between 1986 and 1990. During this period the representation of Canadians of Chinese ancestry was consistently low, whereas in the case of the Canadians of Punjabi ancestry there was improvement approaching equity in the more recent four semesters. Determining whether this improvement will be sustained requires ongoing monitoring. For more specific data on this topic see Beynon et al. (1993).
3. Our first study focused on an analysis of data from questionnaires that asked respondents to rate the relative influence of 12 factors influencing their choice of career. These factors relate to key intrinsic, extrinsic, and interpersonal influences identified by research among the general population into the choice of teaching as a career (Carpenter & Foster, 1977). These factors were: interest in subject matter; opportunity to do work that is personally satisfying; opportunity to do work that helps others (intrinsic); availability of jobs with good salaries; previous employment or volunteer experience (extrinsic); inspiration or encouragement by someone already practicing the career; family members' expectations; parents' wishes for the future; inspiration or encouragement by a teacher or counselor; ethnic group expectations; career choices of friends; and inspiration or encouragement by a guru, minister, rabbi, or priest (interpersonal). Statistically significant differences between ethnic groups occurred only in regard to three interpersonal factors: family members' expectations, parents' wishes and ethnic group expectations. The overall means for these factors were relatively low for all groups (compared with intrinsic and extrinsic factors); but the influence on the minority group individuals was significantly greater than on the Canadians of Anglo-European ancestry.

4. Using 1981 census data Li (1988a) demonstrates that 75% of Canadians of Chinese ancestry were foreign born and over 75% indicated the Chinese language as their mother tongue. This apparent broad similarity does not, however, reveal specific languages or countries of origin.
5. Investigating in detail the similarities as well as contradictions and tensions within and between groups requires historical research, and we know that the two groups focused on here have a daunting variety of historical experiences in immigrating to and residing in this province. This study does not examine many facets of their variation, although we are certain that those differences will be important in characterizing the experiences of individuals and groups.
6. The questionnaire was broadly focused on the myriad factors that could influence career choice: parents, peers, mentors, salaries, critical events, interests, barriers, and so forth. The interviews were guided by open-ended questions about selection of academic program and career. Students were asked to describe their current programs, how they came to be in these, what careers they were planning to go into, the relationship of their family's wishes to their plans. They were asked about their own and their family's and ethnic group's view of being a teacher, and whether there were particular individuals who were influential in guiding and supporting their academic and career choices. In an attempt to ascertain whether their own schooling experiences affected their perceptions of teaching as a career, they were asked whether there were times when they felt all students weren't treated equally, what the reasons for this were, and what if anything teachers could do to meet the needs of students from all groups. Finally, students were asked whether they thought all ethnic groups were equally represented in all careers, what might account for levels of representation (as they saw it) and whether they had any suggestions about ways of increasing the involvement of minorities in teaching.
7. The classification of Canadian occupations used by Li (1988b) and adapted from Wright (1979) is: employers, managers, professionals, petty bourgeoisie and workers.
8. Our findings in the previous study were in concert with Li's (1988 a) analysis of 1981 census data for Canadians of Chinese ancestry showing that about 29% of Canadians of Chinese ancestry had at least some university, compared with about 16% of the rest of the population.
9. Perhaps our study of secondary school students that includes much larger numbers of students who had not yet faced decisions about university attendance would reveal a different picture about the relationship between parent occupation, access to postsecondary education, and desirability of teaching as a career.
10. For students of Chinese ancestry parents' places of birth were almost equally divided between rural China and urban Hong Kong. Eighteen of 22 students identified these as parents' place of birth. However, most of these parents had spent significant parts of their lives in Hong Kong before emigrating to Canada. Of the four remaining students two identified Indonesia, one Vietnam, and one Singapore as parents' (or own) country of origin. All students of Punjabi-Sikh ancestry identified the rural Indian province of Punjab as their parents' place of birth.
11. In a written response to the transcript, Wallace wondered if the major motivation for this incident had been racism or if it had been a matter of older, bigger children bullying a smaller child.
12. Jagdish, who was 33 years old at the time of the interview, was an anomaly in not finding his parents' wishes a major factor in his decision making; we expect that this may have been related to his age.
13. In this study of 50 families only eight were of Sikh background.
14. Gibson and Bhachu (1991) describe a "climate of prejudice permeat[ing] the school experience of all Sikh students" in the farming community in northern California that they studied, such that the Sikh students are "verbally and physically abused by majority students, who refuse to sit with them in class or on buses, crowd in front of them in lines, spit at them, stick them with pins, throw food at them and worse" (p. 69). Like the Canadian students with whom we have been working, these American Sikhs, despite their difficulties, were completing high school, and the boys (at least) were aspiring to postsecondary education.
15. We did not on the basis of our interviews make judgments about language proficiency. As researchers and teachers we are quite accustomed to variations in spoken English and are unlikely to make the same judgments about accent as others might. We are also aware of the research that indicates that oral fluency can be achieved in a significantly shorter period of time (2 to 3 years) than can proficiency in writing (5-7 years) (Cummins, 1986). Thus for the

- purposes of this analysis we only considered the students' self-perceptions of their language proficiency.
16. Punjabi as an Indo-European language is more similar to English than is Chinese and language similarities may well be important in ease and speed of acquisition.
 17. The national controversy about turbans and the RCMP, as well as admission to Canadian Legion halls has put Sikhs in the spotlight in a way that is not apparent in the case of the Chinese. In general, in the Punjabi-Sikh community modes of dress are an important part of cultural and religious practice. For secondary school students where peer dress norms are of tremendous importance, it could be harder for Punjabi than for Chinese students to fit in. The Punjabi experience of ethnic conflict in their country of origin could also be a salient factor in creating sensitivity to and awareness of discrimination in Canada. Although there is certainly a long history of discrimination against both Chinese and Punjabis in Canada, it is not immediately apparent whether most of the Chinese who responded in this study were unaware of or simply unwilling to discuss this matter with our interviewers.
 18. A number of districts in the greater Vancouver area have multicultural/antiracist programs but they are available on a limited basis to selected and self-selected individuals. (Fisher & Echols, 1989). Similarly there are some visible minority teachers in these schools but the students we interviewed had not had any classes with them.
 19. These include incorporation of students' language and culture, collaborative community participation, instructional strategies that draw students into participation, and assessment procedures sensitive to their linguistic and cultural heritage.
 20. Failing provision of these programs, employers can continue to claim that they would hire minority applicants as required by legislation but that there is a dearth of qualified applicants. In the case of educational authorities (school boards and teacher education institutions) it is then possible to resort to the explanation that a particular minority group is simply unsuited to the profession or that parents and students alike are focused on more prestigious and financially rewarding occupations. These are explanations we commonly hear in institutional efforts concerning underrepresentation. We are also aware of the concern many individual minority students have expressed that employment equity legislation singles them out for special treatment simply because of their minority status and not because of their skills and abilities.
 21. Each of these organizations to a greater or lesser extent is doing some work in this area. We do not mean to belittle the importance of the work done by educators committed to transformative education. The problem is that their work is marginalized and encapsulated in a committee, a course, or a department and does not pervade the fundamental operating assumptions of the organization.
 22. A number of the women of Punjabi ancestry noted that there was a Punjabi women's association they belonged to that was very active in mentoring programs, peer counseling, and bridging communication between parents and young women. They indicated that because many young women did not attend temple, this group could be more effective in reaching a wider group of young women.

Acknowledgments

Funding for this study was provided by the SSHRC and by the Faculty of Education and President's Research Grant of Simon Fraser University, Burnaby, British Columbia. We acknowledge Donna Clark, Shemina Hirji, Kauai Keliipio, and Sherry Roberts for their work in interviewing and all the students who generously volunteered to be interviewed. For insightful editorial assistance we thank Dr. Margaret Gold.

References

- Abella, R. (1984). *Equality in employment: A royal commission report*. Ottawa: Minister of Supply and Services Canada.
- Anisef, P., & Okihiro, N. (1982). *Losers and winners*. Toronto, ON: Butterworth.
- Banks, J. (1991). A curriculum for empowerment, action and change. In C. Sleeter (Ed.), *Empowerment through multicultural education* (pp. 125-141). New York: SUNY Press.
- Bannerji, H. (Ed.). (1993). *Returning the gaze: Essays on racism, feminism and politics*. Toronto, ON: Sister Vision Press.
- Baptiste, M.L., Baptiste, H.P., & Gollnick (Eds.). (1980). *Multicultural teacher education: Preparing educators to provide educational equity*. Washington, DC: American Association of Colleges for Teacher Education.

- Bennett, C. (1988). Assessing teachers' abilities for educating multicultural students: The need for conceptual models in teacher education. In C.A. Heid (Ed.), *Multicultural education: Knowledge and perceptions* (pp. 23-38). (ERIC Document Reproduction Service No.ED 312-196)
- Beynon, J., Toohey, K., & Kishor, N. (1993). Do visible minorities in British Columbia want careers in teaching? *Canadian Ethnic Studies*, 24(3), 145-166.
- Beynon, J., & Warsh, M. (1993). Multicultural/anti-racist education: Policy to practice. *Research Forum*, 11, 28-33.
- Breton, R. (1970). Academic stratification in secondary schools and educational plans of students. *Canadian Review of Sociology and Anthropology*, 7, 17-34.
- Buttrick, J. (1977). *Who goes to university from Toronto?* Toronto, ON: Ontario Economic Council.
- Carpenter, P., & Foster, B. (1977). The career decisions of student teachers. *Educational Research and Perspectives*, 4(1), 23-33.
- Carty, L., & Brand, D. (1993). Visible minority women: A creation of the Canadian state. In H. Bannerji (Ed.), *Returning the gaze: Essays on racism, feminism and politics* (pp. 169-182). Toronto, ON: Sister Vision Press.
- Chinn, P., & Wong, G.Y. (1992). Recruiting and retaining Asian/Pacific American teachers. In M.E. Dilworth (Ed.), *Diversity in teacher education: New expectations* (pp. 112-133). San Francisco, CA: Jossey-Bass.
- Clifford, J. (1988). *The predicament of culture: Twentieth century ethnography, literature and art*. Cambridge, MA: Harvard University Press.
- Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review*, 56, 18-36.
- Fisher, D., & Echols, F. (1989). *Evaluation report on the Vancouver School Board's race relations policy*. Vancouver, BC: Vancouver School Board.
- Fleras, A., & Elliot, D.L. (1992). *Multiculturalism in Canada: The challenge of diversity*. Scarborough, ON: Nelson Canada.
- Gibson, M.A., & Bhachu, P.K. (1991). The dynamics of educational decision making: A comparative study of Sikhs in Britain and the United States. In M.A. Gibson & J.U. Ogbu (Eds.), *Minority status and schooling: A comparative study of immigrant and involuntary minorities* (pp. 63-95). New York: Garland.
- Gilbert, S. & McRoberts, H. (1977). Academic stratification and educational plans: A reassessment. *Canadian Review of Sociology and Anthropology*, 14, 34-46.
- Gollnick, D.M., Osayande, K., & Levy, J. (1980). *Multicultural teacher education: Case studies of thirteen programs*. Washington, DC: American Association of Colleges for Teacher Education.
- Government of Canada. (1986). *Employment equity act*. Ottawa: Author.
- Grant, C.A. (Ed.). (1983). Multicultural teacher education—Renewing the discussion: A response to Martin Haberman. *Journal of Teacher Education*, 34, 29-32.
- Hale, S. (1991). Feminist method, process and self-criticism: Interviewing Sudanese women. In S. Berger Gluck & D. Patai (Eds.), *Women's words: The feminist practice of oral history* (pp. 121-136). New York: Routledge.
- Hammersley, M., & Atkinson, P. (1991). *Ethnography: Principles in practice*. London: Routledge.
- Henry, F., & Jain (1991, April 12). When inequality is built right into the system. *The Globe and Mail*.
- Hsia, J. (1988). *Asian Americans in higher education and at work*. Hillsdale, NJ: Erlbaum.
- Le Compte, M.D., & Preissle, J. (1993). *Ethnography and qualitative design in educational research*. San Diego, CA: Academic Press.
- Li, P.S. (1988a). *The Chinese in Canada*. Toronto, ON: Oxford University Press.
- Li, P.S. (1988b). *Ethnic inequality in a class society*. Toronto, ON: Wall and Thompson.
- Lynch, J. (1986). An initial typology of perspectives on staff development for multicultural teacher education. In S. Modgil, G. Verma, K. Mallick, & C. Modgil (Eds.), *Multicultural education: The interminable debate* (pp. 149-165). London: Falmer.
- Martin, R. (1991). The power to empower: Multicultural education for student-teachers. In C. Sleeter (Ed.), *Empowerment through multicultural education* (pp. 287-298). New York: SUNY Press.
- McCarthy, C. (1988). Rethinking liberal and radical perspectives on racial inequality in schooling: Making the case for nonsynchrony. *Harvard Educational Review* 58, 265-279.
- McCarthy, C. (1990). *Race and curriculum*. Bristol, PA: Falmer.
- Ng, R. (1993). Racism, sexism and nation building in Canada. In C. McCarthy & W. Crichlow (Eds.), *Race, identity and representation in education* (pp. 50-59). New York: Routledge.

- Ogbu, J.U. (1991). Immigrant and involuntary minorities in comparative perspective. In M.A. Gibson & J.U. Ogbu (Eds.), *Minority status and schooling* (pp. 3-33). New York: Garland.
- Pang, V.O. (1990). Asian-American children: A diverse population. *Educational Forum*, 55(1), 49-66.
- Pang, V.O. (1991). Teaching children about social issues: Kidpower. In C. Sleeter (Ed.), *Empowerment through multicultural education* (pp. 179-198). New York: SUNY Press.
- Peng, S.S. (1985). *Enrollment patterns of Asian-American students in postsecondary education*. Paper presented at the annual meeting of the American Education Research Association, Chicago.
- Porter, J. (1965). *The vertical mosaic*. Toronto, ON: University of Toronto Press.
- Porter, M., Porter, J., & Blishen, B. (1979). *Does money matter? Prospects of higher education in Ontario*. Toronto, ON: Macmillan.
- Ruiz, R. (1991). The empowerment of language-minority students. In C. Sleeter (Ed.), *Empowerment through multicultural education* (pp. 217-228). New York: SUNY Press.
- Singh Ghuman, P.A. (1993). *Coping with two cultures: British Asian and Indo-Canadian adolescents*. Clevedon, UK: Multilingual Matters.
- Sleeter, C. (1991). Multicultural education and empowerment. In C. Sleeter (Ed.), *Empowerment through multicultural education* (pp. 1-23). New York: SUNY Press.
- Sleeter, C. (1992). *Keepers of the American dream: A study of staff development and multicultural education*. Washington, DC: Falmer.
- Stacey, J. (1991). Can there be a feminist ethnography? In S. Berger Gluck & D. Patai (Eds.), *Women's words: The feminist practice of oral history* (pp. 111-120). New York: Routledge.
- Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. Chicago, IL: University of Chicago Press.
- Vetter, B.M., & Bapco, E.L. (1984). *Professional women and minorities: A manpower data resource service* (5th ed.). Washington, DC: Scientific Manpower Commission.
- Wakil, S.P., Siddique, C.M., & Wakil, F.A. (1981, November). Between two cultures: A study in socialization of children of immigrants. *Journal of Marriage and the Family*, 929-940.
- Wright, E. (1979). *Class structure and income determination*. New York: Academic Press.

Marianne Gareau

and

Don Sawatzky

University of Alberta

Parents and Schools Working Together: A Qualitative Study of Parent-School Collaboration

The term parent-school collaboration is a relatively recent one in education, and little research to date has attempted to study what parents and educators mean when they talk about parent-school collaboration. This article is a report of a study that focused on the perceptions of parents and educators of elementary schoolchildren. In-depth interviews were conducted with five participants: a school principal, a school counselor, a teacher, and two parents. Nine themes emerged from the analysis of these interviews; two were related to the importance of parent-school collaboration, and the remaining seven corresponded to important characteristics of parent-school collaboration. Issues surrounding these themes are discussed.

L'expression collaboration parents-école est tout à fait récente dans le domaine de l'éducation. Peu de recherche a été fait jusqu'à présent pour tenter d'expliquer ce que les enseignant(e)s et les parents veulent dire par la collaboration parents-école. Cet article rapporte une étude qui visait les perceptions des parents et des enseignant(e)s d'enfants scolaires du niveau primaire. Des entrevues détaillées ont été faites avec cinq participant(e)s dont un directeur d'école, un conseiller scolaire, une enseignante et deux parents (femmes). Neuf thèmes sont ressortis de l'analyse de ces entrevues dont deux étaient reliés à l'importance de la collaboration parents-école tandis que les sept autres thèmes correspondaient plutôt aux caractéristiques importantes de la collaboration parents-école. On discute d'autres questions reliées aux thèmes ressortis.

It is generally agreed that good relationships between teachers and parents are important for schools, children, and families. However, the term *parent-school collaboration* is a relatively recent one in education, reflecting a general societal trend toward increased participation of parents in the educational decision-making process. Although the enhancement of parent-school collaboration is often touted as one of the priorities of education departments and school systems, very little research investigates what principals, counselors, teachers, and parents really mean when they talk about such collaboration. Until about 10 years ago, most literature on parent-school relationships focused on the term *parent involvement* (Sattes, 1985). Recently the more frequently used terms are *family-school partnerships* or *parent-school collaboration* (Coleman & Tabin,

Marianne Gareau is a doctoral student in counseling psychology. She has worked as a teacher and consultant with Edmonton Public Schools, and also spent seven years as a consultant for Alberta Education.

Don Sawatzky is a professor in the Department of Educational Psychology and has a special interest in family-school relationships.

1992; Epstein, 1992a). *Collaboration* is generally seen as a broader and more inclusive term than parent involvement because the latter term focuses mainly on the parents' role, whereas the former term focuses on the relationship between the home and the school and how parents and educators work together toward common goals.

The purpose of this study¹ was to understand how parents and educators describe parent-school collaboration, to learn how their views on this topic are similar, and also how they differ from one another. The investigation, conducted as part of the requirements for a master's degree in counseling psychology, employed in-depth, semistructured interviews, and grounded-theory techniques of analysis. The participants for the study included a school principal, a counselor, a teacher, and two parents.

This study was important for a number of reasons. In these times of social and economic challenges and change, it has become more important than ever for educators and parents to find more effective ways to work together for the benefit of children. In generations past, schools were set up as extensions of the family and society; in the past 30 or 40 years we seem to have lost the close relationship that once existed between parents and schools, and in many sectors an adversarial tone seems to have emerged between the school systems and the general public. Walberg (1984) quotes research that shows that the relationship between the home and the school has deteriorated in recent decades. Part of Alberta Education's (1994) Three-Year Business Plan for restructuring education involves increasing parental involvement in education. Another goal listed in the Three-Year Plan involves enabling parents and teachers to have "meaningful roles in decisions about policies, programs, budgets, and activities" (p. 9). In order for these goals to be realized, it is essential to begin by understanding what parents and teachers consider to be meaningful roles.

Current professional literature is filled with the concept of collaboration as a driving theme in education (Cook & Friend, 1991). Collaboration is discussed in terms of relationships in schools and school systems and among schools, families, and communities. Collaboration is defined in the literature in a number of different ways. Appley and Winder (1977) define collaboration as a relational system and suggest that competition and hierarchy no longer serve as an adequate value base for survival. Instead, they believe that we must come to view our environment as one that requires qualitatively new solutions based on an alternative value system, namely, what they call collaboration. In this new value system, caring (as opposed to competition or conflict) is central, and the relational system (rather than the individual) is the basic unit of collaborative effort. In a synthesis of research on organizational collaboration, Hord (1986) reviews the literature and concludes that "while there is little argument about the need for or value of collaboration ... there is disagreement about what 'counts' as collaboration" (p. 22). Hord also suggests that making a distinction between *cooperation* and *collaboration* can be useful because conflicts can arise when it is not clear which model is in process. In cooperation, parties agree only to work together; in collaboration, parties involved are also seen as sharing responsibility and authority for basic decision making. To clarify this distinction between cooperation and collaboration, Hord uses the metaphor of

the family where a mother cooperates with her son by allowing and encouraging his rock band to practice in their home, and the son cooperates with the mother by cleaning the house before guests arrive. On the other hand, the family collaborates in preparing a family meal together, and each offers some form of expertise that is rewarding to all.

Idol and West (1991) define collaboration more as “a structured process and an interactive relationship among individuals” (p. 72). Similarly, Moore and Littlejohn (1992) suggest that collaboration is “an interactive process in which all of the parties are equal” (p. 42), emphasizing the importance of equality of parents and educators. On the other hand, Cook and Friend (1991) refer to collaboration as a style rather than a process.

Historical and Theoretical Perspectives of School-family Relationships

There is no doubt that parents and families have always been involved in the education of their children, to the point that most writers refer to parents as being the oldest and most essential part of any education system (Berger, 1987). Parent involvement in the education of children has in fact been present since prehistoric times, with the family providing most of children’s education informally until the introduction of public schooling. In the early 1900s the parent education movement began, and both parent education and involvement became institutionalized parts of the schools through parent-teacher associations. Until the 1940s, parent involvement in the schools, though indirect, was meaningful and constant (Comer, 1986). In small towns and rural areas parents had everyday contact with teachers. There was a sense of community and cultural uniformity, and the schools were natural extensions of the community. After World War II, however, came many technological, scientific, cultural, and social changes that resulted in a change in the level of trust and agreement between home and school. Training and achievement requirements were raised for educators and students alike; the content of the curriculum began to change, along with increased focus on teaching methods. By the 1950s the major responsibility for children’s formal, academic education was delegated almost entirely to the schools (Topping, 1986). Schools discouraged parental intrusion, and the educative function of parents was downplayed.

Gradually the pendulum swung back once again, and the trend changed toward increased parent involvement in the schools. Comer and Haynes (1991) point out that the involvement of parents in their children’s education is now widely accepted as desirable and even essential to effective schooling. However, they say that a significant number of educators are ambivalent about parent involvement in the schools, and parent participation is not significant in many schools even when parents are invited. Moles (1987), while promoting parent participation in education as an idea whose time has come, points out that the elements of confrontation and power sharing tend to make educators and school officials uneasy with parents in the roles of advocates and decision makers, roles where parents are becoming increasingly interested. Studies indicate (Lightfoot, 1978; Ost, 1988) that the majority of teachers and principals see the ideal relationship with parents as one in which parents support teacher practices and schools in general, carry out requests, but do not interfere with plans and decisions. Although parent involvement in instruction has been clearly linked to student success (Epstein, 1992a), teachers and administrators

often fail to establish strong links between home and school (Collinge & Coleman, 1992).

Recent Research on Parent-school Collaboration

A 1988 study by Leitch and Tangri on barriers to home-school collaboration focused on junior high schools and assessed both parents' and schools' concerns. The most frequent barrier as seen by the teachers was the parents, and their attitudes toward the school. When parents were asked to identify barriers to collaborating with the schools, they also centered the barriers in themselves, citing work responsibilities, lack of time, health problems, and economic differences between themselves and teachers. From this study, Leitch and Tangri (1988) concluded that the major barriers to collaboration are the lack of specific planning and the lack of knowledge about how teachers and parents can use each other more effectively.

Recent studies by Lindle (1989) and Lindle and Boyd (1991) asked parents to reflect on the worst and the best experiences they had with any schools and also asked teachers what they thought parent preferences were. Surprisingly, the preferences of parents were not what school personnel thought they were; according to this study parents viewed a professional, businesslike manner as undesirable and reported a personal touch as the most enhancing factor in school relations. All parents seemed to prefer relationships of a less formal nature with their child's teachers. Lindle (1989) concluded that parents do not want a professional-client relationship with the school: they want to be seen as equal partners with schools in the rearing of children.

Another group of researchers at Simon Fraser University have developed an intervention study that attempts to identify the attitudes among parents, teachers, and students that influence collaboration. They conclude that it must be teachers who permit parent-school collaboration (Coleman & Tabin, 1992); that is, they must realize the importance of teacher invitation, legitimizing, facilitating, and encouraging parent collaboration as part of their roles as teachers. However, Coleman and Tabin (1992) find little evidence that teachers are presently doing this; most classrooms are not characterized by strong positive relationships with parents. Coleman, Collinge, and Seifert (1992) reject the concept of parents as clients of the school, which they believe implies an arm's-length relationship. They also see the notion of parents as clients as inherently at odds with the notion of collaboration, which they believe is the more desirable relationship between schools and parents. Instead, they describe parents and students as participants in the school community and believe that "parents will have more influence, and will be more satisfied with the schools, if they are perceived as partners rather than as consumers or adversaries" (Coleman et al., 1992).

There is a definite trend in our society toward recognizing the rights of parents to make educational decisions concerning their children, along with a trend that emphasizes the importance of parents and educators working together to meet the educational needs of the children in their care. The present study was undertaken in an attempt to understand how parents and educators in Alberta describe the current reality in terms of parent-school collaboration. Furthermore, we wanted to learn more about the participants' hopes or vision for the future in terms of the relationship between parents and the schools.

Methodology

We chose a qualitative, human science research approach (Bogdan & Biklen, 1992; Patton, 1990; van Manen, 1990), because of its congruence with the topic and our personal beliefs about how we can best learn about another person's perspectives. In making a decision regarding the participants in this study, we were aware that the most basic qualification was that participants have "salient experiences of the phenomenon in their everyday worlds" (Becker, 1986, p. 105). Each of the participants in this study had direct experience of parent-school collaboration. Because one of the goals of the study was to explore the diversity of various perspectives on parent-school collaboration, we chose to interview a principal, a teacher, a counselor, and two parents. The decision to include two parents was made in order to provide some balance between the school/parent perspectives. Finally, we decided to select all the participants from one school in the hope that there would be some common ground in their experiences. Each of the participants was either a staff member of Holy Spirit Elementary School or had children who attended the school. Following is a brief description of each of the participants (names of individuals and the school have been changed to keep identities anonymous).

Stance, a principal. Stance has been the principal at Holy Spirit School for the past five years. He is married, has two children, and has been working as a teacher and administrator for over 20 years.

Bernard, a counselor. Bernard has been the school counselor at Holy Spirit for 17 years and has been working for the school district for 23 years. He is also married and has three children.

Clare, a grade 3 teacher who has been teaching elementary school for 20 years. She has also worked as a district office consultant for two years. She is married to a high school teacher and has three daughters.

Anne, a parent of two children (Julie, grade 5 and Justin, grade 3). Anne has a son in Clare's class this year, and Clare suggested her as a parent participant for the study. Anne is a nurse and a single parent and has raised both children on her own since they were very young.

Denae, a parent of three children (Veronique, grade 5, Melanie, grade 4, and Lisa, grade 1). Denae's name was suggested for this study by Bernard. Denae is a full-time homemaker, and her husband is a self-employed businessman.

Individual, in-depth semistructured interviews were used as the primary basis for the study in order to obtain as rich data as possible (Patton, 1990). An interview guide was followed, but participants were encouraged to digress and expand on their thoughts, and probing questions were used to elicit greater depth of responses. The opening item for each interview was: "Tell me about parents and schools working together." Other questions included asking participants how they knew when parent-school collaboration was occurring, asking them to describe a situation where they felt collaboration *had* occurred and another situation where it *had not* occurred, asking them to identify barriers to collaboration, to define or describe collaboration, and to tell about their vision of collaboration. I found that one of the later questions in the interview, "Is there anything else that I need to know about you personally in order to really understand your thoughts about this topic?" was the question that seemed to yield some of the richest data.

Each participant was interviewed twice, the first an open-ended interview, and the second a validation of my interpretations. Our intention during the interviews was to be as nondirective as possible, to avoid influencing the participants' responses. The first participant to be interviewed was Stance, the principal; followed by Bernard, the counselor; Clare, the teacher; and finally, Anne and Denae, the parents. With the exception of Stance, who was interviewed in late February 1994, all other participants were first interviewed during a two-week period in May 1994.

The data for the study consisted of the five taped interviews, the written transcriptions of these interviews, field notes, and journal reflections. In analyzing the data, each of the interview transcripts was first broken down into meaning units, and each unit was carefully paraphrased (first-level abstraction). Working with the help of a colleague, we then assigned a tag or label to each meaning unit (second-level abstraction). These tags were then clustered together into groups or topics for analysis, in order to develop the main themes for each participant (third-level abstraction). At this point, we returned to each participant for the second interview, in order to validate our analysis by seeing if each one agreed with it. After analyzing each interview individually, we began a comparative analysis. First we looked for themes that were specific to parents and specific to educators; second, we looked for some of the overall themes or issues (topics) that applied to all the interviews.

Findings

Nine major themes or issues emerged from the analysis of the interviews. The first two themes (children live in two worlds, and changes in society) address the question of the importance of parent-school collaboration, whereas the other seven themes (communication, trust/openness/honesty, positive/caring attitudes, personal connections, being equals, power/conflict/roles, and schoolwide commitment) can be seen as characteristics of that collaboration.

Why is parent-school collaboration so important?

Children live in two worlds

Each of the participants in the study was committed to the value and importance of parent-school collaboration. The most frequently mentioned reason for this was related to the "diffuse boundaries" for the child between the home and the school, or what Christenson, Rounds and Franklin (1992) refer to as the "mutually influencing quality" among home and school experiences. For example, Anne described parent-school collaboration as an ongoing relationship, where both parties have to work as partners because they both have a vested interest in the child:

Not that they're the educators and I'm the parent, and when he comes home from school, that their job ends and mine begins, it's an ongoing relationship that happens, from day to day.

Clare, the teacher, also talked about being keenly aware of the connection between the home and the school for the child:

I keep seeing, like, it isn't the child here, totally distinct, in this classroom, from his home life ... It's not, we carry them back and forth, the child carries [his

experiences] back and forth ... We need to realize that the connection is so important.

Changes in society

The other theme that emerged when participants spoke about the importance of parent-school collaboration was that of change. Without exception, everyone spoke about his or her belief that we are in a time of change with respect to the roles of parents and schools. Some of the participants saw this change as connected with the current political situation in Alberta, with budgetary restrictions that may result in expecting parents to play a greater role in the governance and implementation of educational programs. Participants also spoke about how this generation of parents has more education, as well as higher expectations of the schools—they are “involved more, questioning more.” Bernard, the counselor, talked about how he believes that collaboration between schools and parents has pretty much been kept at a “surface level” until quite recently: “We’re only beginning to see what it really means for parents and teachers and other school personnel to really work closely together.” He sees that parents are now more knowledgeable and are less willing to just accept the schools having total authority over educational decisions. Stance, the principal, also saw himself in process and changing some of his own views about the role of parents. He said that he realized that parents have not yet been brought in as full members of the educational team.

What are the characteristics of collaboration?

Communication

Both Clare, the teacher, and Anne, a parent, spoke about collaboration mainly in terms of the communication between the home and the school. All the participants believed that without communication, collaboration was impossible. However, for the parents and the teacher, communication was seen mainly in an individual, child-related sense; for Stance, the principal, a major concern was that parents and teachers communicate openly and honestly, as well as listening to various points of view. For both parents “being informed,” “feeling listened to,” and “having a voice” were among their strongest needs in their relationship with the school.

Trust, openness, honesty

Trust was talked about as an important characteristic in parent-school collaboration by all the participants. However, not everyone means the same thing when they talk about trust. For example, for both parents trusting the school meant knowing that the teachers would be able and willing to look after their children’s needs and recognize each child’s uniqueness. For Denae trust also meant that the school would keep her informed about what was going on and that it would not do things that were incompatible with the school’s Christian mission. On the other hand, when the educators spoke about trust they were referring more to the importance of parents and educators having a basic belief in each other’s good intention. For example, Stance would like parents to trust that the school is doing their best and that the educators are “professionally to know” what is in the child’s best educational interest.

Each of the participants spoke about how trust and the mutual respect that ensues are developed gradually over time. There needs to be an openness and

honesty between people in order for this trust to develop; participants all believed that we can only reach the deeper levels of collaboration once these are established.

Positive, supportive, caring attitudes

Everyone spoke about the importance of parents and educators being positive and supportive toward one another. Clare talked about how, for her, collaboration has to be based on an environment of caring; Denae insisted that her comments must be viewed as supportive and positive, and not interpreted as critical. Stance said that he found that negative people “really get in the way of collaboration,” and Bernard said that people must give each other positive recognition for what each has to offer.

Personal connections

In Lindle’s (1989) research, establishing personal connections was seen by parents to be the most important characteristic of a collaborative relationship. This finding was also recognized as significant by both parents and educators in the present study. For Denae, knowing that the school people “really care” for the children is the most important thing. Stance talked about many of the informal contacts that he makes with parents, such as calling them at home to share some of the good things that their children have done at school. Clare, the teacher, establishes a relationship on a first-name basis with parents, makes face-to-face contact whenever possible, and talks about the bond that she feels with the parents of her students. Bernard, as the counselor, also emphasized the importance of caring for the parents as well as the students. In talking about this personal connection, the educators in this study felt that they were the ones with the greater responsibility for initiating this type of relationship.

Being equals

Each of the participants spoke about the importance of parents and educators seeing each other as equals in order for true collaboration to occur. The parents did not feel that this sense of equality always existed, although they recognized and appreciated it when it did. Bernard said he believed that “if we don’t begin with a strong commitment to equality between parents and the school, the whole aspect of team is gone.” For Stance, one of the ways for this to happen is for parents and educators to begin to “share roles a bit,” for educators to give up some of their “professionalism” and to acknowledge that parents also have a lot to teach them and to contribute to the child’s education.

Power, conflict, and roles

Differences in power between parents and educators is an issue that seems to be of concern mainly to the parents. Both parents talked of being worried about how the educators were perceiving them. They feel vulnerable in a way that seems qualitatively different for parents than it does for the educators. As Anne told me about some of her interactions with the school concerning her son Justin’s program, she said, “There were times when I thought they were thinking, ‘here she comes again!’” Similarly, Denae worried that Stance might misinterpret her statements at the parent meeting where she disagreed with him, and thought that he might think that she was not being supportive of him and of the school. Parents often feel powerless when they interact with school

personnel. Denae summed up some of her feelings by saying, "It's like it's a partnership as long as we all agree."

For the educators the concern is more related to what Clare refers to as the fine line regarding parents' power to make school decisions. She expresses concern about how much ownership parents should have for education. Stance also shared this concern, saying that he is "never sure at what point it's appropriate to make sure that there are safeguards in place." Stance says that although he recognizes that parents and educators share responsibilities, there are still some areas where the lines have been drawn, at least in an informal way. Stance admits that he finds this confusing and says that he "wrestles all the time" with trying to determine what the areas are where parents might have as "full a range of decision making" as teachers and administrators. Recognizing that we are in a time of changing roles and expectations, he says, "I'm just not sure where the parameters are right now, anymore."

Although everyone in this study recognized that conflict is inevitable, not everyone is comfortable in its presence. Denae said that she feels very uncomfortable with conflict and dissension, feeling that "dissension does not serve any purpose." Bernard, however, believes that if people are truly working together, conflict is inevitable, and that unless we begin to address the conflict areas, we will remain at what he calls a surface level in terms of collaboration. In talking about conflict Stance explained his belief that there needs to be at least a minimum level of basic agreement before collaboration between parents and the school can occur.

Schoolwide commitment

The importance of having a schoolwide commitment to collaborative relationships is one that all three educators in this study mentioned. As they said, collaboration must start at the school level; it is highly doubtful that educators who are not engaged in collaborative relationships among themselves will engage in collaborative relationships with parents. At Holy Spirit School there appears to be an atmosphere of working together that has also had an impact on parent-school relationships. As Clare said, "collaboration is not something that can happen in isolation; it needs to be a schoolwide approach to working with others." Stance believes that it is essential that all the movement toward collaboration not come from the administrators; he encourages parents to deal directly with their children's teachers whenever possible. The parents in the study also recognized the atmosphere that is established in the school, an "upbeat, positive one that extends to all the families that come to the school." All of the participants mentioned the crucial role that the administrators have in establishing this schoolwide commitment to collaborative relationships.

An Important Dilemma

One of the important dilemmas highlighted by this research is related to the difficulty of introducing a collaborative culture into what has traditionally been a hierarchical organization. It appears that school personnel experience feelings of vulnerability related to the potential loss of some of their professionalism. As long as school personnel relate the role of professional to that of expert, it will be difficult to achieve more than surface-level collaboration. In the way that both parents and educators referred to the issue of trust, for

example, it became evident that the school was strongly viewed as the body that would ultimately make educational decisions.

The focus of the parents in the study was to be able to trust the school to make these important educational decisions concerning their child. Both parents shared feelings of vulnerability vis-à-vis the school and concerns about how they might be viewed by the school and how this might in turn affect their children. This feeling of vulnerability was not expressed by the educators in the study, although they all shared an interest in hearing from parents in honest and open exchanges. All three educators in this study saw the ideal relationship with parents as going beyond parents being merely supportive of the school. At the same time, they expressed concerns about how far they could go in sharing their power with parents before they lost some essential autonomy as professionals. Educators here seem to be looking for some balance (Litwak & Meyer, 1974) and in this sense are sharing some of the concerns that emerge as the roles of parents in the education process are reexamined. Although we as educators verbalize definitions of collaboration, there are indications that we are still falling short of understanding and implementing true collaboration. The term collaboration has been discussed here in terms of the introduction of a new culture based on equal participation of parents and educators in educational decisions.

The dilemma exists for both parents and educators. For parents it entails relinquishing the security that comes when someone else has final responsibility for making educational decisions. For educators, what has to be relinquished is the notion of professionalism that is based on hierarchy and power.

The Role of the School Counselor

Epstein (1992b) notes that few leaders in the field of school-family relations have been school psychologists or counselors, in spite of the natural connections between their skills and interests and the needs of schools for better school and family connections. Over 10 years ago Lombana and Lombana (1982) suggested that if parent-school partnerships are to be strengthened in the face of opposing cultural, geographic, and attitudinal forces, school counselors should assume leadership roles in the school, particularly in developing and implementing plans for working with parents. They also stressed that these must be plans that meet the expressed parental needs, as well as ensure the most productive use of available resources. Noting that teachers and administrators are generally ill-prepared to work with parents, Lombana and Lombana (1982) advocated for counselors to take a more active role in the school in consulting with and supporting teachers in their work with parents.

These ideas proposed by Lombana and Lombana (1982) are closely related to those of Weiss and Edwards (1992), who suggest that school psychologists might assume the role of family-school coordinators; and Epstein (1992b) who advocates for the school psychologist to become the psychologist of the school. There have been some longstanding debates about ways for school psychologists to broaden their influence by working with all staff, students, and families (Epstein, 1992b) and to extend some of the special education practices to all families in a school. The results of this study would support these ideas and also give some direction to school psychologists who are interested in broadening their roles.

Conclusions

The term *parent-school collaboration* is a relatively recent one in education, reflecting a significant societal paradigm shift that is also occurring in other institutions such as health care and social services. Parents and schools have shared responsibility for children's education ever since public schools came into existence, but the recognition that parents might have equal rights in educational decision making is a concept that has emerged only over the past two or three decades (Cross, 1989). It is becoming an increasingly accepted fact that education, like other social services in our democratic society, is everybody's business; it is especially the business of those parents who have children in schools.

Note

1. This study was part of a master's thesis written by the first author.

References

- Alberta Education. (1994). *Meeting the challenge: Three-year business plan*. Edmonton, AB: Author.
- Appley, D.G., & Winder, A.E. (1977). An evolving definition of collaboration and some implications for the world of work. *Journal of Applied Behavioral Science*, 13, 279-290.
- Becker, C.S. (1986). Interviewing in human science research. *Methods*, 1, 101-124.
- Berger, E.H. (1987). *Parents as partners in education: The school and home working together* (2nd ed.). Toronto, ON: Merrill.
- Bogdan, R.C., & Biklen, S.K. (1992). *Qualitative research for education*. Toronto, ON: Allyn and Bacon.
- Christenson, S.L., Rounds, T., & Franklin, M.J. (1992). Home-school collaboration: Effects, issues, and opportunities. In S.L. Christenson & J.C. Conoley (Eds.), *Home-school collaboration: Enhancing children's academic and social competence*. Silver Spring, MD: National Association of School Psychologists.
- Coleman, P., Collinge, J., & Seifert, T. (1992). *Seeking the levers of change: Participant attitudes and school improvement*. Paper presented at the Fifth Annual International Congress on School Effectiveness and Improvement, Victoria.
- Coleman, P., & Tabin, Y. (1992). *The good teacher: A parent perspective*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Collinge, J., & Coleman, P. (1992). *Home-school collaboration: Teacher practices and parent attitudes*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Comer, J.P. (1986). Parent participation in the schools. *Phi Delta Kappan*, 67, 442-446.
- Comer, J.P., & Haynes, H.M. (1991). Parent involvement in schools: An ecological approach. *Elementary School Journal*, 91(3), 271-277.
- Cook, L., & Friend, M. (1991). Principles for the practice of collaboration in schools. *Preventing School Failure*, 35(4), 6-9.
- Cross, J. (1989). Recording: Parent, professionals and their perceptions of partnership. *Scottish Educational Review*, 21(2), 106-116.
- Epstein, J.L. (1992a). School and family partnerships. In M.C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1139-1151), sponsored by the American Educational Research Association. New York: Macmillan.
- Epstein, J.L. (1992b). School and family partnerships: Leadership roles for school psychologists. In S.L. Christenson & J.C. Conoley (Eds.), *Home-school collaboration: Enhancing children's academic and social competence* (pp. 499-515). Silver Spring, MD: National Association of School Psychologists.
- Hord, S.M. (1986). A synthesis of research on organizational collaboration. *Educational Leadership*, 43(5), 22-26.
- Idol, L., & West, J.F. (1991). Educational collaboration: A catalyst for effective schooling. *Intervention*, 27(2), 70-78.
- Leitch, M.L., & Tangri, S.S. (1988). Barriers to home-school collaboration. *Educational Horizons*, 66(2), 70-74.
- Lightfoot, S.L. (1978). *Worlds apart: Relationships between families and schools*. New York: Basic Books.

- Lindle, J.C. (1989). What do parents want from principals and teachers? *Educational Leadership*, 47(2), 12-14.
- Lindle, J., & Boyd, W.L. (1991). Parents, professionalism, and partnership in school-community relations. *International Journal of Educational Research*, 15(3/4), 323-337.
- Litwak, E., & Meyer, H.J. (1974). *School, family and neighborhood: The theory and practice of school-community relations*. New York: Columbia University Press.
- Lombana, J.H., & Lombana, A.E. (1982). The home-school partnership: A model for counselors. *Personnel and Guidance Journal*, 61(1), 35-39.
- Moles, O.C. (1987). Who wants parent involvement? Interest, skills, and opportunities among parents and educators. *Education and Urban Society*, 19(2), 137-145.
- Moore, D., & Littlejohn, W. (1992). Trends toward increased family-school collaboration. *Contemporary Education*, 64(1), 40-45.
- Ost, D.H. (1988). Teacher-parent interactions: An effective school-community environment. *Educational Forum*, 52(2), 163-176.
- Patton, M.Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Sattes, B.D. (1985). Parent involvement: A review of the literature. *Appalachia Educational Laboratory Occasional Paper Series*. Charleston, WV: Educational Services Office, Appalachia Educational Laboratory.
- Topping, K. (1986). *Parents as educators: Training parents to teach their children*. London: Croom Helm.
- van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. London, ON: University of Western Ontario.
- Walberg, H.J. (1984). Families as partners in educational productivity. *Phi Delta Kappan*, 65, 397-400.
- Weiss, H.M., & Edwards, M.E. (1992). The family-school collaboration project: Systemic interventions for school improvement. In S.L. Christenson & J.C. Conoley (Eds.), *Home-school collaboration: Enhancing children's academic and social competence*. Silver Spring, MD: National Association of School Psychologists.

G. Patrick O'Neill
Brock University

Authorship Patterns in AJER: Forty Years in the Making

This article examines authorship patterns in AJER from 1955 to 1994. Data were collected on the number of single, co- and multiple authorships for each year. The years were, in turn, divided into eight five-year intervals. Frequencies and percentages were generated for each interval and chi-squares were computed between intervals and overall. In addition, author/article ratios were calculated for each year. Results showed increases in the number of co- and multiple authorships with corresponding decreases in the number of single authorships. However, the trend was broken in 1990. The author/article ratios fluctuated somewhat, but generally paralleled the shifts within intervals. It was tentatively concluded that education is not unlike other fields, that it too has experienced changes in authorship patterns. Interpretations are offered along with recommendations for further research.

Cet article examine le nombre d'auteur(e)s qui contribuèrent à la revue AJER entre 1955 et 1994. Les données ont été collectionnées pour chaque année en utilisant les articles écrits par un(e) auteur(e), par deux auteur(e)s, et par multiples auteur(e)s. On divisa ensuite les années en huit intervalles de cinq ans chacune. Les fréquences et les pourcentages ont été produits pour chaque intervalle et des χ^2 ont été calculés entre ces intervalles et sur la période entière. En plus, la proportion d'auteur(e)/article a été calculée pour chaque année. Les résultats indiquent une croissance d'articles écrits par deux ou multiples auteur(e)s et une décroissance d'articles écrits par un(e) auteur(e) seulement. Cependant, l'année 1990 dévie de cette tendance. Les proportions auteur(e)/article fluctuaient quelque peu mais elles égalaient généralement les changements à l'intérieur des intervalles. On peut conclure de façon tentative que l'éducation ne diffère pas des autres domaines et qu'en éducation même retrouve-t-on des changements dans les nombres d'auteur(e)s qui écrivent des articles. Cet article offre des interprétations ainsi que des recommandations pour une recherche plus poussée.

Introduction

Recently authorship patterns in journals have been explored in relation to de Solla Price's (1963) famous three predictions. He posited that single authorships would be extinct by 1980, that more than half of all articles published by 1980 would be multiple-authored, and that we would "move steadily toward an infinity of authors per paper" (p. 89). His predictions have been tracked in fields such as anthropology (Choi, 1988), business and economics (Petty & Kerr, 1982), counseling (Gladding, 1984; Strahan, 1982; Zook II, 1987), finance (Schweser, 1983), life sciences (Burman, 1982; Satyanarayana & Ratnakar, 1989), medicine (Burman, 1982; de Villiers, 1984; Farr, 1984; Powers, 1988; Rosenfeld, 1991; Satyanarayana & Ratnakar, 1989), nursing (Norris, 1993), physical education (Cruse & Rasato, 1992), and psychology (Mendenhall & Higbee, 1982; Sacco & Milana, 1984). The studies themselves have varied in

G. Patrick O'Neill is a professor in the Graduate Department, Faculty of Education. His interests are in the areas of special education, teacher education, and higher education.

design, but all have come to the same conclusion. In short, there has been a significant increase over the years in coauthored (i.e., double, dual, joint) or multiple (i.e., three or more) authored articles, or both. The increases, though, have not always supported de Solla Price's (1963) predictions. For example, not one of the studies supports his first prediction; some, particularly in medicine, support his second prediction; and not one supports his third prediction.

Background

To date there have been only two authorship studies in education, neither of which has been Canadian. The first investigation (Lyle, 1986) looked at authorship patterns in volumes one through 22 (1963 to 1984) of the *Journal of Industrial Teacher Education*. He excluded editorials, book reviews, and other special sections from his analysis. Lyle found that the number of single authorships dropped below 70% in 1971 and have remained so for all but two years since. In contrast, the number of coauthorships increased from 7% in 1964 to a high of 44% in 1981. As well, coauthorships accounted for more than 20% of all authorships between 1971 and 1984. Multiple authorships showed a similar pattern, but remained substantially below that of joint authorships. They ranged from zero to a high of 16%. The author/article ratios reflected these changes being generally higher for the years 1971 to 1984.

The second inquiry (Swafford, 1990) examined authorship patterns in the *Journal of Educational Administration* from 1963 to 1987. All articles and research reports were included in the analysis. Excluded were editorials and book reviews. In order to detect trends, Swafford divided his analysis into five equal time periods as follows: 1963-1967, 1968-1972, 1973-1977, 1978-1982, and 1983-1987. He found that between 1963 and 1967 only 2% of the articles were coauthored. However, this figure rose to 18.2% for the 1968-1972 period and to 29.6% for the 1973-1977 period and has remained fairly steady since. Multiple authorships showed no discernible pattern as they accounted for fewer than 3% of all articles published between 1963 and 1987.

Problem

The purpose of this study is to look at authorship patterns in a Canadian education journal, namely, the *Alberta Journal of Educational Research (AJER)*. Specifically, this study addresses the following six questions:

1. Overall, have coauthorships increased in *AJER*?
2. Overall, have multiple authorships increased in *AJER*?
3. If so, what pattern, if any, do the increases follow?
4. Is there a statistically significant relationship between the number of authors per article and dates of publication?
5. Overall, has there been an increase in *AJER*'s author/article ratio?
6. If so, what pattern, if any, do the ratios follow?

Procedures

AJER was chosen as the target journal because it is the oldest national refereed journal in education and one of the most highly respected in the country. All major articles including perspectives and symposia from issue 1, volume 1, 1955 to issue 4, volume 40, 1994, were included in the analysis. Excluded were editorials, book reviews, essay reviews, journal reviews, rejoinders and replies, research abstracts and notes, memorials, bibliographies, department reports,

book notices, conference notices, and advertisements. The data were grouped into eight five-year intervals as follows: 1955-1959, 1960-1964, 1965-1969, 1970-1974, 1975-1979, 1980-1984, 1985-1989, and 1990-1994. Each interval contained 20 issues.

The analysis was divided into two parts. The first addressed questions one to four; the second addressed questions five and six. Part one included the compilation of a record sheet for each of the eight intervals. There were three steps. First, the number of authors, that is, single, double, and multiple for each year and for the total interval was tabulated. Second, the number of authors per category was converted to percentages for each year and for the total interval. Third, chi-squares were calculated between each interval and overall. A 2 × 2 design was employed in which single authorships comprised one category and co- and multiple authorships comprised the other.

The second part of the analysis included the calculation of 40 author/article ratios. This was achieved by dividing the number of articles per year by the number of authors per year.

Findings

Table 1 summarizes a typical record sheet that was compiled for each interval. Included is the number of *AJER* authors per article for the years 1955 to 1959 inclusive. As illustrated, the period was dominated by single authorships. More than 85% of the authorships were single whereas joint authorships accounted for about 13% and multiple authorships accounted for about 2% of the total.

Table 2 gives the breakdown for the eight five-year intervals. Included are the totals for each period. As shown, single authorships accounted for more than 58% of all authorships between 1960 and 1974. However, the figure dropped to less than 50% between 1975 and 1989. At the same time, there was a corresponding rise in both joint and multiple authorships. In fact dual and multiple authorships accounted for more than 50% of all articles between 1975 and 1989. The biggest increase was in two-authored articles. Multiple authorships remained consistently below that of coauthorships, although they too increased sharply between 1980 and 1989.

Table 1
Percentage and Number of *AJER* Authors per Article for the Years 1955-1959

No. of Authors	1955		1956		1957		1958		1959		Totals	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
One	68.42	(13)	95.24	(20)	90.90	(20)	86.96	(20)	83.33	(20)	85.32	(93)
Two	26.32	(5)	4.76	(1)	4.55	(1)	13.04	(3)	16.67	(4)	12.85	(14)
Three or more	5.26	(1)	—	(0)	4.55	(1)	—	(0)	—	(0)	1.83	(2)
Totals	100.00	(19)	100.00	(21)	100.00	(22)	100.00	(23)	100.00	(24)	100.00	(109)

Table 2
Percentage and Number of AJER Authors per Article for the Eight Intervals Between 1955 and 1994

No. of authors	1955-1959		1960-1964		1965-1969		1970-1974		1975-1979		1980-1984		1985-1989		1990-1994	
	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n
One	85.32	(93)	66.47	(84)	58.65	(78)	60.90	(95)	49.30	(71)	42.65	(58)	46.72	(57)	57.66	(79)
Two	12.85	(14)	33.33	(42)	35.34	(47)	30.77	(48)	44.45	(64)	42.65	(58)	38.52	(47)	29.20	(40)
Three or more	1.83	(2)	—	(0)	6.01	(8)	8.33	(13)	6.25	(9)	14.70	(20)	14.76	(18)	13.14	(18)
Totals	100	(109)	100	(126)	100	(133)	100	(156)	100	(144)	100	(136)	100	(122)	100	(137)
χ^2		6.64 ³		1.78		.51		4.07 ²		1.25		.43		3.10 ¹		33.86 ⁴

¹p<.10
²p<.05
³p<.01
⁴p<.001

The latest interval (1990-1994) proved to be somewhat of an anomaly. The number of multiple authorships remained fairly constant, but there was a shift in the number of double and single authorships.

The chi-square values between intervals and overall are given at the bottom of Table 2. The chi-square values were significant between the first two intervals ($\chi^2 (1, N = 235) = 6.65, p < .01$), the fourth and fifth intervals ($\chi^2 (1, N = 300) = 4.07, p < .05$), the seventh and eighth intervals ($\chi^2 (1, N = 259) = 3.10, p < .10$), and overall ($\chi^2 (1, N = 1063) = 33.86, p < .001$). These data would suggest that there is indeed a significant relationship between the number of authors per article and dates of publication. That is, the more recent the publication, the greater the number of authors. And although the relationship is not always significant, the overall is significant indicating that the direction is continuous, at least between 1955 and 1989.

Table 3 presents the author/article ratio for the years 1955 to 1994. The ratios fluctuated somewhat, but generally show marked increases, especially in 1960, 1965, 1968, 1970, 1974, 1980, 1984, 1986, and 1992. The lowest ratio was in 1956 (1.05); the highest was in 1992 (2.13). This means that there were 1.05 authors per article in 1956 compared with 2.13 authors per article in 1992.

Discussion

The findings of this inquiry are by and large similar to those of Lyle (1986) and Swafford (1990). All three studies found substantial increases in the number of co- and multiple authored articles although the numbers overall in *AJER* were somewhat higher. These results, then, would suggest that authorship differences, if any, between education and other fields are in degree, not pattern.

Evidence would also indicate that there may be plateaus in *AJER*'s authorship patterns. Plateaus are defined as, "spans of time in which the number of single, double and multiple authorships remains relatively stable" (Norris, 1993, p. 156). Accordingly, then, 1960 to 1974 and 1975 to 1989 would be two successive plateaus, each in turn having increased, but fairly consistent numbers of dual and multiple authored articles. This observation is supported in part by the fact that the chi-squares were significant at the beginning and end, but not within each of the periods.

If the trend were to continue, 1990 would be the beginning of another 15-year plateau, one where single authorships would again show marked decreases and joint and multiple authorships would in turn show marked increases. But this does not happen. Instead, there is a sudden reversal in the number of single and double authorships. And, if it were not for 1992, which had a disproportionately high number of dual authorships, the reversal overall would be far more pronounced.

This swing in direction would suggest that we may be entering an unexpected plateau, one where *AJER* is once more dominated by single authorships. Then, of course, the shift may be temporary, an aberration. In any case it is too early to predict the next plateau with any degree of accuracy because present data are limited to only five years and there are no comparison data. Future studies might address both these limitations.

Table 3
AJER Author/Article Ratio for the Years 1955-1994

Year	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964
Articles	19	21	22	23	24	24	23	24	28	27
Authors	26	22	25	26	28	35	30	30	35	38
Ratio	1.37	1.05	1.14	1.13	1.17	1.46	1.30	1.25	1.25	1.41

Year	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974
Articles	30	25	33	24	21	24	28	35	34	35
Authors	48	35	45	38	31	40	42	48	47	58
Ratio	1.60	1.40	1.36	1.58	1.48	1.67	1.50	1.37	1.38	1.66

Year	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
Articles	29	35	31	24	25	21	33	31	26	25
Authors	43	57	52	38	37	38	55	53	41	49
Ratio	1.48	1.63	1.68	1.58	1.48	1.81	1.67	1.71	1.58	1.96

Year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Articles	23	25	21	29	24	27	24	23	33	30
Authors	35	50	33	49	41	37	35	49	57	48
Ratio	1.52	2.00	1.57	1.69	1.71	1.37	1.46	2.13	1.73	1.63

Interpretation

Reasons for the increase in double and multiple authorships in *AJER* are speculative, but nonetheless worthy of note. I can think of at least three explanations; others may occur to the reader. First is the ever-increasing pressure on faculty to "publish or perish" (Day, 1994; Gladding, 1984; Mendenhall & Higbee, 1982; Sacco & Milana, 1984; Strahan, 1982; Zook II, 1987). This combined with an institutional reward structure that values single and joint authorships equally, as many do (Pettry & Kerr, 1982), no doubt, helps account for the increased number of coauthorships.

Second, and closely allied, is the increase in collaborative efforts by graduate students and their supervisors. Indeed, the practice of mentoring has become more common (Mendenhall & Higbee, 1982; Zook II, 1987) for without "an established record of publication, Ph.D. holders will have little opportunity to obtain a teaching or research position in higher education" (Dill & Morrison, 1985, p. 177). I can identify with this factor, in particular, having coauthored an article in *AJER* with my advisor while pursuing a doctorate.

Third, perhaps, is the increasing complexity of research itself (Mendenhall & Higbee, 1982; Powers, 1988; Sacco & Milana, 1984; Zook II, 1987). The knowledge explosion combined with increased specialization and large, often interdisciplinary projects have forced researchers to pool their resources. In other words,

As experimental methods have become more complex, and as complex statistical analyses have warranted a knowledge of computer applications, the need and likelihood for researchers to work together has increased. Researchers have divided the work according to each author's area of research expertise (e.g., subject matter, computer knowledge, methodology, statistics). (Mendenhall & Higbee, 1982, p. 6)

Conclusion

Clearly these results do not support de Solla Price's (1963) predictions. The number of authors per article has increased, but not to the extent proposed. For instance, the three-or-more category did not exceed 15% of all authorships in any one interval. This is a far cry from the 50% plus predicted by de Solla Price. In addition, there is no sign of a movement "toward an infinity of authors per paper." In fact, in the 40-year history of the journal, only four articles had five or more authors. The first appeared in volume 20, issue 3, 1974 and has six authors. The second appeared in volume 32, issue 1, 1986 and has five authors. The third and fourth articles appeared in volume 38, issue 1, 1992. The first has five authors; the second seven authors (the record holder).

Moreover, there may be a movement back to single authorships. Should this occur, we might safely conclude that de Solla Price's predictions were myopic; that, in reality, authorship patterns are cyclic, not linear.

References

- Burman, K.D. (1982). Hanging from the masthead: Reflections on authorship. *Annals of Internal Medicine*, 97, 602-605.
- Choi, J.M. (1988). An analysis of authorship in anthropology journals, 1963 and 1983. *Behavioral and Social Sciences Librarian*, 6, 85-94.
- Cruse, D., & Rascato, F.D. (1992). Single versus multiple authorship in professional journals. *Journal of Physical Education, Recreation and Dance*, 63, 28-31.
- Day, R.A. (1994). *How to write and publish a scientific paper*. Phoenix, AZ: Oryx Press.

- de Solla Price, D.J. (1963). *Little science big science*. New York: Columbia University Press.
- de Villiers, F.P. (1984). Publish or perish—The growing trend towards multiple authorship. *SA Medical Journal*, 66, 882-883.
- Dill, D.D., & Morrison, J.L. (1985). Ed.D. and Ph.D. research training in the field of higher education: A survey and a proposal. *Review of Higher Education*, 8, 169-186.
- Farr, A.D. (1984). Multiple authorship of scientific papers. *Medical Laboratory Sciences*, 41, 1-2.
- Gladding, S.T. (1984). Multiple authorship in the *Personnel and Guidance Journal*: A 12-year study. *Personnel and Guidance Journal*, 62, 628-630.
- Lyle, S.P. (1986). Multiple authorship in volumes 1 through 22 of the *Journal of Industrial Teacher Education*. *Journal of Industrial Teacher Education*, 24, 105-108.
- Mendenhall, M., & Higbee, K.L. (1982). *Recent trends in multiple authorship in psychology*. Paper presented at the Annual Meeting of the Western Psychological Association, Sacramento. (ERIC Document Reproduction Service No. ED 222 807)
- Norris, R.P. (1993). Authorship patterns in CJNR: 1970-1991. *Scientometrics*, 28, 152-158.
- Petry, G.H., & Kerr, H.S. (1982). Pressure to publish increases incidence of co-authorship. *Phi Delta Kappan*, 63, 495.
- Powers, R.D. (1988). Multiple authorship, basic research, and other trends in the emergency medicine literature (1975 to 1986). *American Journal of Emergency Medicine*, 6, 647-650.
- Rosenfeld, R.M. (1991). Clinical research in otolaryngology journals. *Archives of Otolaryngology—Head and Neck Surgery*, 117, 164-170.
- Sacco, W.P., & Milana, S. (1984). Increase in number of authors per article in ten APA journals: 1960-1980. *Cognitive Therapy and Research*, 8, 77-84.
- Satyanarayana, K., & Ratnakar, K.V. (1989). Authorship patterns in life sciences, preclinical basic and clinical research papers. *Scientometrics*, 17, 363-371.
- Schweser, C. (1983). The economics of academic publishing. *Journal of Economic Education*, 14, 60-64.
- Strahan, R.R. (1982). More on JCP publication: Single versus multiple authorship. *Journal of Counselling Psychology*, 29, 430-431.
- Swafford, G.L. (1990). Window or mirror? A content analysis of the first 25 years of the *Journal of Educational Administration*. *Journal of Educational Administration*, 28, 5-23.
- Zook II, A. (1987). Trend toward multiple authorship: Update and extension. *Journal of Counselling Psychology*, 34, 77-79.

R.S. Brown

J.O. Anderson
University of Victoria.

and

M.M. McGee

Manifestation of Latent Thetas: A Comparison of Field-Study and Positivistic Approaches to the Investigation of Self-Nurturance

Self-nurturance is an essential trait of humans that appears to have its genesis in survival motivation (Gecko, 1985). It is thought that a balanced manifestation of self-nurturance leads to an independent yet meaningful interactive life in the human community (Rondolini, 1988a, 1988b, 1994). However, there is an absolute dearth of empirical studies into the phenomenon.

This article reports on the findings generated by a study into manifest self-nurturance. Two research approaches were used to collect and analyze data. A field study approach was used to observe and interview individuals who were self-described self-nurturants. The researchers shared the lives of a small, selected sample of individuals in order to explore the meanings and manifestations of self-nurturance. A positivistic approach was taken in another component of the study. This involved surveying a larger, cross-cultural sample of individuals and subjecting the data to analyses based on bivariate individual score curve theory (Rogers, McLean, & Hambleton, 1992). A final comparative analysis was planned in which the model generated from the field based study was tested using the empirical survey data in a procrustean, structural equation modeling approach.

The Field Study

The field study involved a prolonged, participatory observation of a purposely selected dyad of self-identified self-nurturants who had been living in Victoria, British Columbia for over a decade. This identification of self-nurturance was supported by both anecdotal descriptions generated from a conversation between researchers and the individuals of the dyad, and categorization of life-elements on the G-scale by a panel of trained raters (Gucci, 1982). The basic approach to data collection adopted was a multi-staged infiltration protocol (Brown, 1989). The initial engagement began with the senior author sharing the daily lives of the dyad. To minimize intrusive characteristics of the protocol, he

R.S. Brown and M.M. McGee are independent field researchers living in Victoria, British Columbia.

J.O. Anderson teaches in the Faculty of Education at the University of Victoria.

deliberately role-played a submissive in the social setting of the home. During this period, blatant recording of observations such as field notes, audio- or videotape could be not used unobtrusively; rather, recollective anecdotes were the prime data archived from the study. The participant observation continued over a three-year period and succeeded to the extent that the researcher effectively became a member of a triad of self-nurturants.

Concerns about the fidelity of the data coupled with the apparent success of the infiltration prompted the insertion of a second researcher into this social universe of self-nurturance. The junior author joined the ethnography in the third year at the invitation extended by the first researcher but supported by the original dyad. The staged infiltration resulted in two levels of anecdotal evidence allowing for a rich information base that could be fully investigated for strong patterns and evaluated by means of fold-over credence methodology (Traub, 1993).

The information collected over what has turned out to be a five-year period provides significant and consistent patterns of insight into the meaning and manifestations of self-nurturance that are reported in detail elsewhere (Anderson & Brown, 1994). It should be pointed out that the information collected suffered from what could best be termed restricted range. Given the social and psychological situation of the observed dyad and the reluctance of the researchers to impose direction on the inclinations and behaviors of the observed dyad, what has been collected provides an insider perspective on what could best be termed the indulgent extreme of self-nurturance. No one, not the observeds nor the observers, was willing to explore the lived experience of denial and privation. However, given the data that was collected (essentially the mid to extreme range of self-nurturance) projections can be made to the nihilist extreme.

For this article the three major themes are briefly described: bliss, rationalization, and the golden extreme. *Bliss* is the prime target state, the state of mind and soul of the individual after a period of growth and development induced by self-nurturant behaviors. A classic example of observed bliss occurred annually in the study dyad. Generally in mid-February when the cherry trees were blooming (the study was located in Victoria) the obvious self-nurturant behaviors of both individuals of the dyad included repeated comments regarding the "absolute beauty of the flowers and their elegant simplicity as individual flowers ... and voluptuous complexity when viewed in magnitude."

The dyad increased the frequency of watching the weather channel noting the temperatures and conditions on the prairies and *down east*. This frequency peaked when snowstorms and extremely low temperatures persisted, and this was accompanied by telephone calls to folks back home to see how things were. This was self-nurturant in the sense of contributing to the growth of satisfaction and adaptation to their life condition, thereby stabilizing the social, psychic, and physiological entropy of existence, in this way tending toward a state of bliss. The state of bliss was often accompanied by what some may consider trite comments, but to the individuals in a state of bliss were accurate, and perhaps profound, reflections of the phenomenon: "My, but isn't life grand" or "Ahh, this is wonderful."

Some (Maguire, 1976; Allard, 1985) have used the alternate term *smug* to label this particular manifestation, but we deemed that it did not capture the positive valencing attributed the state by the individuals who live the experience.

Rationalization was another strong theme in the information. Individuals often expressed reasons as to why self-nurturant behaviors were of merit. For example, an affinity for red wines of one of the participants was described in a growth-related manner in regard to the increased perceptivity and creativeness caused by the consumption of the beverage, and further, that recent articles in the popular press attributed enhanced health and longevity to the consumption of red wine. This was viewed as not only of personal benefit, but of general societal benefit not only in the contribution to the coffers of the federal treasury through exorbitant taxation on the alcohol, but also in the reduced drain on the medicare system over the long term. Such protracted, convoluted, and yet genuinely believed positions were held for most self-nurturant behaviors recognized by the individuals of the dyad.

The golden extreme was a pattern derived from the observational data rather than any text collected from conversation. If the self-nurturance involved consumption, red wine consumption for example, moderation would occur only if some external environmental constraints were imposed such as a lack of funds to obtain sufficient quantities. Otherwise consumption would be maximized. Another example occurred with the use of meditation that was viewed as self-nurturant in that mental restoration and invigoration, reestablishment of psychic center, and general rebalancing of the self could be attained. Both individuals practiced meditation to the point of prolonged catharsis, and in fact preferred it as a complete lifestyle. The only reason they maintained active professional lives was the need for revenue to provide what they viewed as necessary resources.

The Cross-cultural Survey

A survey instrument developed in a previous study (McGee, 1979) to measure the levels of general life satisfaction and self-nurturance (GLSSN) was administered to samples of individuals from the West Coast, the Prairies, central Ontario, Quebec, and the Maritimes. These populations were chosen to maximize variation. The GLSSN is a forced-choice, polychotomously scored, computer-mediated instrument that can be administered either individually using the specially developed GLSSN headgear or in groups by means of a Jumbotron. Demographics (home location, sex, age, and height) were also collected.

The GLSSN yields two scores for each individual: General Life Satisfaction (GLS) and Level of Self-Nurturance (LSN). Analysis included calculation of descriptive statistics and correlations and bivariate individual score curve theory analysis (BISCT). BISCT (Rogers et al., 1992) is essentially an inverted item response theory (IRT: Lord, 1980) approach. IRT characterizes item performance on the basis of person responses, whereas BISCT characterizes people on the basis of test scores. BISCT generates theta values that are descriptive of the bivariate status of an individual on the full range of the satisfaction/self-nurturant plane.

The first round of analysis was conducted on each sample separately and results were compared across samples. No significant differences were detected, so for further analyses data were pooled.

Descriptive statistics of the GLSSN were best summarized as a frequency distribution for the total sample of respondents (Figure 1). The Self-Nurturant scores were remetricated utilizing the nosological standardization algorithm of Fossey (1948) thus resulting in a scale of 0.0 to 1.0. From this distribution it can be seen that self-nurturance is essentially normally distributed within the pooled data. The average value was 0.62 with a standard deviation of 0.28. The extremes of the distribution were labeled deprivation and indulgence for what should be obvious reasons, the center could best be termed balanced self-nurturance (BSN).

The distribution of these data was not unexpected given the recent findings and opinions of Klien (1993). However, the empirical nature of these data certainly lends credence to the notion of the inherent ability of Canadians to look after themselves. Of greater import and uniqueness are the results of the BISCT analyses that yielded Self-Nurturant/Satisfaction plots (SNSPs). From these plots it became evident that within the pool of respondents were archetypical profiles that were more interpretable than simple distributions of pooled data, and were more amenable to the planned procrustean merger (Heitz, 1987) with the observation data of the field studies. Each of the found patterns is based on data from at least 300 respondents, and the resulting characteristic person curve (CPC) was smoothed using the LOWESS procedures (Wilkinson, 1990). It was believed that these merged data would then lead to the development of the structural model of self-nurturance.

The SNSP's resulted in highly interpretable representations of the data that resonated with the observed life of self-nurturants investigated in the field study component of these researches. On the basis of the research literature, a number of major patterns had been anticipated and were indeed found. The first major pattern anticipated was what we termed the *golden mean* (Figure 2a),

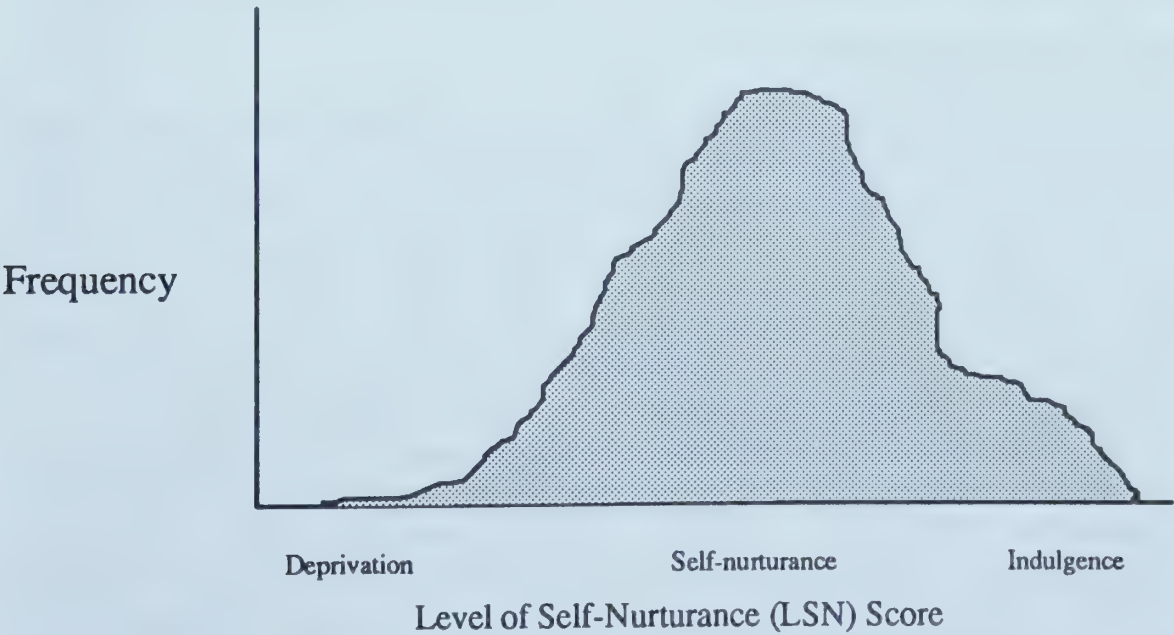


Figure 1. Distribution of self-nurturance scores in total sample of respondents (n=7,877).

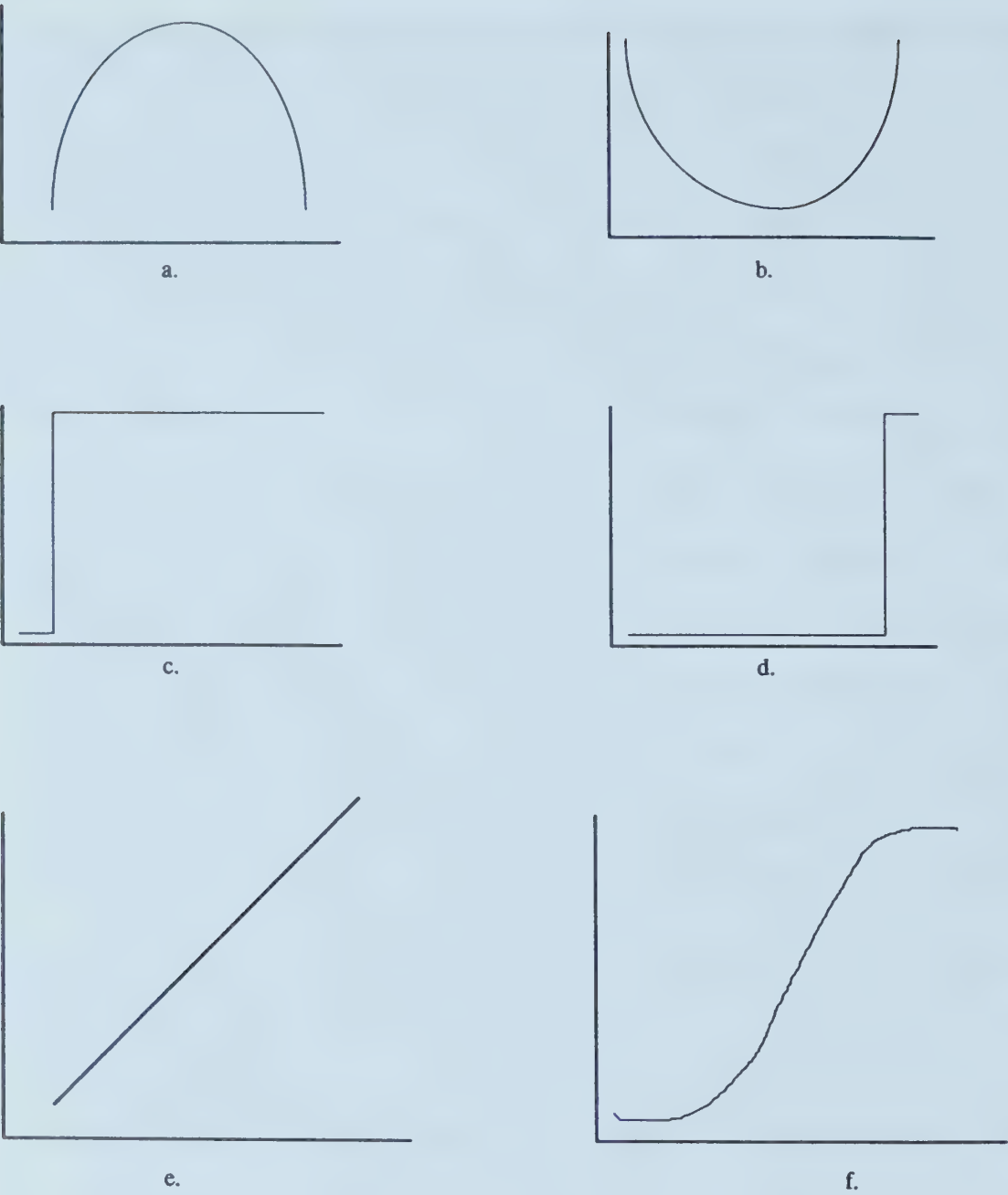


Figure 2. Person characteristic curves.

too much (or little) of a good thing is not all that satisfying. Rather, just enough is just right. Another pattern that was initially believed to be an artifact of data inversion but on further analysis proved to be accurate was the *guilt-ridden-masochistic indulgent* (Figure 2b) who is characterized as being most satisfied at the extremes of self-nurturance. From some of the available observational information, the basis to the lack of satisfaction with moderate levels of self-nurturance is associated with voluminous guilt that apparently dissipates as the state of indulgence is approached or an aesthetic privation is attained.

Departures from the continuous relationships to a dichotomous condition were also found (predominantly with data generated from administrators and accountants). The profile we labeled the *pollyanna* (Figure 2c) indicates an individual who is maximally satisfied with minimal levels of self-nurturance. The *sulky hedonist* (Figure 2d) on the other hand, was maximally dissatisfied with all but the indulgent levels of self-nurturance. Incidentally, the dichots

were generally reported to be extremely difficult individuals to live with given that they were either insufferably satisfied or unsatisfied.

The pattern associated with the majority of respondents, particularly those from the ranks of academe, shows a constantly increasing satisfaction with increasing levels of self-nurturance, the *material realist* (Figure 2e). The material realist simply is more satisfied with more self-nurturance, apparently without limit according to our data. A variation on this pattern is a *modulated material realist* (Figure 2f), which is characterized by limits to the change in satisfaction at the extremes of self-nurturance. The modulation is more characteristic of southern Ontario, whereas the nonmodulated material realist is more common among residents of the West Coast.

Some exploratory analyses conducted on small ($n < 100$) datasets revealed some potentially significant variations. Figure 3 demonstrates what are termed guilt-dips in the otherwise linear relationship of satisfaction to self-nurturance in the material realist. It appears that some mr's actually have episodic lapses in increasing satisfaction as self-nurturance increases, a pattern we are confident is not associated with anyone reading this article.

Conclusion

It is obvious that the findings of both the field study and the survey allow for profound understandings of the human condition. It was the plan to synthesize the two datasets to develop and test a structural model of self-nurturance, and some work has proceeded in this direction. However, as this article was in preparation the field team defected to their study dyad, in effect creating a self-indulgent quadrad that has rejected the notion of research as a valued activity. In doing so, they absconded with all existing field notes and associated data. This results in an abandonment of the third stage of the project: the procrustean fitting of the empirical GLSSN data with the model of self-nurturance developed from the insights and meanings generated from the dyad of the field study. A preliminary model and novel display protocol had been developed but not yet fully evaluated. It is a sad note on which to end but it does accurately reflect the dangers of delving into the heart of human self-nurturance.



Figure 3. Person characteristic curves with guilt-dips.

References

- Allard, M. (1985). *Bliss my soul what's wrong with me?* Memphis, TN: Hound Dog Press.
- Anderson, J.O., & Brown, R.S. (1994). *Tending the personal garden*. Toronto, ON: Harrowsmith.
- Brown, R.S. (1989). Aiming for non-obtrusiveness in observation. *Methodology in Observation*, 25(4), 233-235.
- Fossey, D. (1948). Cross-specific standardization in primates. *Agronomist*, 46(1), 23-27.
- Gecko, G. (1985). *Greed is good*. New York: Simian Press
- Gucci, A. (1982). *Evaluating nurturance of individuals and groups*. Milan: Moderna Press.
- Heitz, H.H. (1987). Blending qualitative and quantitative data meaningfully. *Journal of Applied Qualitative Analysis*, 2(2), 134-149.
- Klien, R. (1993). The suppression of self-nurturance in academe. *Western Reports*, 11(6), 344-348.
- Lord, F. (1980). *Applications of item response theory to practical testing problems*. New York: Erhlbaum.
- Maguire, T.O. (1983). *Individualism as random rogue variables*. Millet, AB: Populist Press.
- McGee, Q. (1979). Two essential dimensions of life: Satisfaction and self-nurturance. *WestCoast Life*, 12(3), 125-133.
- Rogers, W.T., McLean, L.D., & Hambleton, R. (1992). *Individual score theory applications*. Denver, CO: Mile High Books.
- Rondolini, H.L. (1988a). Self-nurturance and student achievement. *Journal of Educational Measurement*, 24(3), 345-356.
- Rondolini, H.L. (1988b). Predictors of optimal self-nurturance: A factor analytic approach. *Multivariate Behavioral Psychology*, 12(4), 545-560.
- Rondolini, H.L. (1994). Meanings of self-nurturance in the developing self. *Annals of Lived Hermeneutic Interpretations*, 3(1), 33-103.
- Traub, R. (1993). Fold-over credence checks in multi-level information sets. *Journal of Applied Qualitative Analysis*, 7(2), 234-239.
- Wilkinson, L. (1990). SYSTAT. Evanston, IL: SYSTAT Inc.

Index, Volume XLI, 1995

- Amadeo Marquez, S. *The Spurious Relationship Between IQ and Social Behavior: Ethnic Abuse, Gender Ignorance, and Confounded Education*, No. 3, 277-282.
- Anderson, J.O. *Manifestation of Latent Thetas: A Comparison of Field-Study and Positivistic Approaches to the Investigation of Self-Nurturance*, No. 4, 482-488.
- Andrews, J.J.W. *A Classroom-based Social Skills Intervention for Children with Learning Disabilities*, No. 1, 84-102.
- Barnett-Foster, D. *A Comparison of Undergraduate Test Response Strategies for Multiple-choice and Constructed-response Questions*, No. 1, 18-35.
- Barrow, R. *Keep them Bells a-Tolling*, No. 3, 289-296.
- Bateson, D.J. *How Can The Bell Curve Be Taken Seriously?* No. 3, 271-273.
- Belke, T.W. *A Synopsis of Herrnstein and Murray's The Bell Curve: Intelligence and Class Structure in American Life* (New York, Free Press, 1994), No. 3, 238-256.
- Beynon, J. *Access and Aspirations: Careers in Teaching as seen by Canadian University Students of Chinese and Punjabi-Sikh Ancestry*, No. 4, 435-361.
- Biggs, J. *Assessing for Learning: Some Dimensions Underlying New Approaches to Educational Assessment*, No. 1, 1-17.
- Bountrogianni, M. *Teachers' Views of Parents: Family Decision Making Styles and Teacher-Parent Agreement Regarding Homework Practices and Values*, No. 2, 175-187.
- Brannigan, A. *IQ and Crime: Dull Behavior and/or Misspecified Theory?* No. 3, 316-321.
- Brown, R.S. *Manifestation of Latent Thetas: A Comparison of Field-Study and Positivistic Approaches to the Investigation of Self-Nurturance*, No. 4, 482-488.
- Cameron, J. *Introduction to Special Issue: Canadian Perspectives of The Bell Curve*, No 3, 237-237.
- Cameron, J. *From the Editor*, No. 4, 377-377.
- Connors, J.B. *Two Tails are Better than One: The Logic of The Bell Curve*, No. 3, 342-348.
- Conte, R. *A Classroom-based Social Skills Intervention for Children with Learning Disabilities*, No. 1, 84-102.
- Cousins, J.B. *Giving and Receiving Explanations in Cooperative Learning Groups*, No. 1, 103-121.
- da Costa, J.L. *Teacher Collaboration: A Comparison of Four Strategies*, No. 4, 407-420.
- Derwing, T. *Review of Genie: An Abused Child's Flight from Silence* by Russ Rymer, No. 1, 122-124.
- Dolmage, R.W. *Accusations of Teacher Sexual Abuse of Students in Ontario Schools: Some Preliminary Findings*, No. 2, 127-144.
- Epling, W.F. *The Bell Bottom*, No. 3, 302-303.
- Fancher, R.E. *The Bell Curve on Separated Twins*, No. 3, 265-270.
- Filipovich, T. *Teachers' Views of Parents: Family Decision Making Styles and Teacher-Parent Agreement Regarding Homework Practices and Values*, No. 2, 175-187.
- Gambell, T.J. *Ethnography as Veneration*, No. 2, 162-174.
- Gareau, M. *Parents and Schools Working Together: A Qualitative Study of Parent-School Collaboration*, No. 4, 462-473.
- Gartrell, J. *The Spurious Relationship Between IQ and Social Behavior: Ethnic Abuse, Gender Ignorance, and Confounded Education*, No. 3, 277-282.

- Gartrell, J. *University Performance of Nonmatriculated Admissions*, No. 4, 421-434.
- Gibbs Van Brunschot, E. *IQ and Crime: Dull Behavior and/or Misspecified Theory?* No. 3, 316-321.
- Golec, J. *University Performance of Nonmatriculated Admissions*, No. 4, 421-434.
- Greene, M.L. *The Forgotten Link: Transitions from Graduate School to Classroom Teaching*, No. 2, 213-230.
- Holdaway, E.A. *Expectations Held for Department Heads in Postsecondary Institutions*, No. 2, 188-212.
- Hull, J. *Indian Control and the Delivery of Special Education Services to Students in Band-operated Schools in Manitoba*, No. 1, 36-62.
- Hutton, G. *A Classroom-based Social Skills Intervention for Children with Learning Disabilities*, No. 1, 84-102.
- Jones, D.R. *Expectations Held for Department Heads in Postsecondary Institutions*, No. 2, 188-212.
- Kavanagh, M. *The Bell Curve: Rising Fears, Fallen Dreams*, No. 3, 297-301.
- Kirby, J.R. *Intelligence and Social Policy*, No. 3, 322-334.
- Krishnan, P. *The Bell Curve: Some Statistical Concerns*, No. 3, 274-276.
- Krull, C.D. *IQ Testing in America: A Victim of its Own Success*, No. 3, 349-354.
- Livingstone, D.W. *For Whom The Bell Curve Toils*, No. 3, 335-341.
- Loomer, M. *A Classroom-based Social Skills Intervention for Children with Learning Disabilities*, No. 1, 84-102.
- MacPherson, E.D. *The Time Machine: The Next Generation*, No. 3, 304-307.
- Malicky, G.V. *Perceptions of Literacy and Adult Literacy Programs*, No. 1, 63-83.
- McGee, M.M. *Manifestation of Latent Thetas: A Comparison of Field-Study and Positivistic Approaches to the Investigation of Self-Nurturance*, No. 4, 482-488.
- Nagy, P. *A Comparison of Undergraduate Test Response Strategies for Multiple-choice and Constructed-response Questions*, No. 1, 18-35.
- Norman, C.A. *Perceptions of Literacy and Adult Literacy Programs*, No. 1, 63-83.
- Norris, S.P. *Format Effects on Critical Thinking Test Performance*, No. 4, 378-406.
- O'Neill, G.P. *Authorship Patterns in AJER: Forty Years in the Making*, No. 4, 474-481.
- Phillips, R. *Indian Control and the Delivery of Special Education Services to Students in Band-operated Schools in Manitoba*, No. 1, 36-62.
- Pierce, W.D. *IQ Testing in America: A Victim of its Own Success*, No. 3, 349-354.
- Polyzoi, E. *Indian Control and the Delivery of Special Education Services to Students in Band-operated Schools in Manitoba*, No. 1, 36-62.
- Pratt, M.W. *Teachers' Views of Parents: Family Decision Making Styles and Teacher-Parent Agreement Regarding Homework Practices and Values*, No. 2, 175-187.
- Priegert Coulter, R.P. *And Academic Sexism Too: A Comment on The Bell Curve*, No. 3, 308-311.
- Purvis, C.R. *The Forgotten Link: Transitions from Graduate School to Classroom Teaching*, No. 2, 213-230.
- Reviere, R. *"Slow Through The Dark": A Commentary on The Bell Curve: Intelligence and Class Structure in American Life*, No. 3, 312-315.
- Robeck, E.C. *Distorted Vision: Education As Seen Through The Bell Curve*, No. 3, 367-376.
- Robitaille, D.F. *Distorted Vision: Education As Seen Through The Bell Curve*, No. 3, 367-376.

- Ross, J.A. *Giving and Receiving Explanations in Cooperative Learning Groups*, No. 1, 103-121.
- Sawatzky, D. *Parents and Schools Working Together: A Qualitative Study of Parent-School Collaboration*, No. 4, 462-473.
- Schleifer, M. *Should we Change our Views About Early Childhood Education?* No. 3, 355-359.
- Siegel, L.S. *Does the IQ God Exist?* No. 3, 283-288.
- Stenlund, K.V. *Teacher Perceptions Across Cultures: The Impact of Students on Teacher Enthusiasm and Discouragement in a Cross-cultural Context*, No. 2, 145-161.
- Sveinson, W.J. *University Performance of Nonmatriculated Admissions*, No. 4, 421-434.
- Symaluk, D.G. *Book Review: Women's Work is Never Done: Resolving Gender Inequalities in Education*, No. 2, 231-235.
- Toohy, K. *Access and Aspirations: Careers in Teaching as seen by Canadian University Students of Chinese and Punjabi-Sikh Ancestry*, No. 4, 435-361.
- Wahlsten, D. *Increasing the Raw Intelligence of a Nation is Constrained by Ignorance, Not its Citizens' Genes*, No. 3, 257-264.
- Wangler, D.G. *Is The Bell Curve A Ringer?* 360-366.
- Wolbring, G. *Review of Violence and Abuse in the Lives of People With Disabilities* by Dick Sobsey, No. 1, 124-126.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format	Manuscript in triplicate. Typewritten or printed, double-spaced throughout. Abstract of approximately 100 words on a separate sheet. Manuscript not to exceed 6,500 words excluding graphics. Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.
Referencing	Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s) for direct quotes. Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format. Page numbers must be given. All sources listed alphabetically at the end of the manuscript under the heading <i>References</i> using the APA style. Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading <i>Notes</i> ; any citations in notes follow the same format as other references.
Graphics	Each table and figure on a separate page at the end of the manuscript. Figures, charts, and diagrams in camera-ready form. Number tables in Arabic numerals and center over each heading, e.g., <i>Table 1</i> . Number figures in Arabic numerals before the caption at foot of each graph, e.g., <i>Figure 1. ...</i> Indicate placement of figures and tables in text, i.e., <i>Insert Table 3 here</i> .
Style	AJER's editorial style conforms closely to the <i>Publication Manual of the American Psychological Association</i> . Editorial changes may be made to manuscripts. For spelling, consult <i>Webster's New Collegiate Dictionary</i> ; spelling in quoted material remains as in the original.
Diskettes	If manuscript is accepted for publication please supply a diskette from a Macintosh, IBM, or IBM compatible system. Microsoft Word and WordPerfect are the preferred software.

Submission to AJER implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index*, *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *ERIC/Current Index to Journals in Education*, and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts*, *Educational Administration Abstracts*, *Educational Technology Abstracts*, *Family Abstracts*, *Language Behavior Abstracts*, *Multicultural Education Abstracts*, *Psychological Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Sociology of Education Abstracts*, and reviewed by the *Journal for Research in Mathematics Education*.

If undelivered return to:

The Alberta Journal of Educational Research
845 Education Centre South
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1438